DESCRIPTION

Title of Invention

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

Technical Field

[0001] An aspect of the present invention relates to an image predictive encoding device, image predictive encoding method, and image predictive encoding program. Another aspect of the present invention relates to an image predictive decoding device, image predictive decoding method, and image predictive decoding program. Particularly, these aspects relate to an 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 to generate a predicted signal of a target block using motion information of surrounding blocks. Still another aspect of the present invention relates to a video encoding device, video encoding method, video encoding program, video decoding device, video decoding method, and video decoding program to generate a predicted signal with motion compensation by a motion vector.

Background Art

[0002] The compression encoding technologies are used for efficient

transmission and storage of still images and video data. The systems of MPEG-1 to 4 and ITU (International Telecommunication Union) H.261 to H.264 are commonly used for video data.

[0003] In these encoding systems, an image serving as an encoding target is partitioned into a plurality of blocks and then an encoding process or a decoding process is carried out. In intra-picture predictive encoding, a predicted signal is generated using a neighboring previously-reconstructed image signal (i.e. signal reconstructed from compressed image data) present in the same picture as a target block and then a differential signal obtained by subtracting the predicted signal from a signal of the target block is encoded. In inter-picture predictive encoding, motion compensation is performed with reference to a neighboring previously-reconstructed image signal present in a picture different from a target block to generate a predicted signal, and a difference signal obtained by subtracting the predicted signal from 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-reconstructed pixel values neighboring a block serving as an encoding target, in a predetermined direction. Fig. 22 is a schematic diagram for explaining the intra-picture prediction method used in ITU H.264. In (A) of Fig. 22, target block 802 is a block serving as an encoding target, and a pixel group 801 consisting of pixels P_A-P_L neighboring a boundary of the target block 802 is a neighboring region, which is an image signal previously reconstructed in past processing.

[0005] In the case shown in (A) of Fig. 22, the predicted signal is

generated by downwardly extending the pixel group 801 which includes neighboring pixels located immediately above the target block 802. In the case shown in (B) of Fig. 22, the predicted signal is generated by rightwardly extending previously-reconstructed pixels (P_I-P_L) located to the left of the target block 804. Specific methods for generation of the predicted signal are described, for example, in Patent Literature 1. A difference is calculated between each of nine predicted signals generated by the methods shown in (A) to (I) of Fig. 22 in the above-described manner, and the pixel signal of the target block, and the predicted signal with the smallest difference is selected 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 typical inter-picture predictive encoding, the predicted signal for a block serving as an encoding target is generated by a method of searching previously-reconstructed pictures for a signal similar to a pixel signal of the target block. A motion vector that is a spatial displacement amount between the target block and a region composed of the detected signal, and a residual signal between the pixel signal of the target block and the predicted signal are then encoded. The searching technique for a motion vector for each block as described above is called block matching.

[0007] Fig. 21 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 target block 702 on encoding target picture 701. Picture 703 is a previously-reconstructed picture and region 704 is a

region located spatially at the same position as target block 702. In the block matching, a search range 705 around the region 704 is set and a region 706 to minimize the sum of absolute differences from the pixel signal of the target block 702 is detected from a pixel signal of this search range. The signal of this region 706 is determined to be a predicted signal, and a displacement amount from region 704 to region 706 is detected as motion vector 707. Also employed is a method of preparing a plurality of reference pictures 703, selecting a reference picture to be used in block matching for each target block, and detecting reference picture selection information. In H.264 are a plurality of prediction types of different block sizes for encoding of motion vectors are prepared, in order to adapt for local feature changes of images. 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 about the encoding order in the inter-picture prediction to generate the predicted signal with reference to previously-reconstructed pictures. The first technique is forward prediction to generate the predicted signal with reference to a past previously-reconstructed picture in a reproduction order, the second technique is backward prediction to generate the predicted signal with reference to a future previously-reconstructed picture in the reproduction 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 HEVC (High efficiency video coding) under standardization as a next-generation video encoding system, introduction of asymmetric partitions as shown in (E) to (F) of Fig. 20 is also in review, in addition to rectangular bisections shown in (B) and (C) of Fig. 20 and square quartering shown in (D) of Fig. 20, as partitioning types of a prediction block. In HEVC, a further technique in review is to use motion information (the motion vector, the reference picture information, and the inter-picture prediction mode to identify the forward/backward/bidirectional prediction) of a block neighboring a target prediction block serving as a prediction target, on the occasion of generating the predicted signal of the prediction block partitioned in this manner. This prediction technique is called block merging and is characterized by enabling efficient coding of motion information. (A) of Fig. 2 is a drawing schematically illustrating neighboring blocks in block merging with prediction block T1 generated by vertical partition of encoding block 400. The predicted signal of prediction block T1 is generated using 1) motion information of neighboring block A, 2) motion information of neighboring block B, or 3) motion information detected in block matching. For example, when an encoder selects the motion information of neighboring block A, the encoder first sets merge identification information (merge_flag) indicative of use of motion information of neighboring block to "merge_flag=1" and transmits the merge identification information (merge_flag) to a decoder. Second, the encoder sets merge block selection information (merge_flag_left) indicative of use of neighboring block A out of neighboring block A and

neighboring block B, to "merge_flag_left=1" and transmits the merge block selection information (merge_flag_left) to the decoder. The decoder, receiving the two pieces of information, can identify that the predicted signal of the target prediction block should be generated using the motion information of neighboring block A. Similarly, when the decoder receives "merge_flag=1" and "merge_flag_left=0" (selection of neighboring block B), it can identify that the predicted signal of the target prediction block should be generated using the motion information of neighboring block B; when it receives "merge_flag=0," it can identify that it should further receive the motion information from the encoder, and restore the motion information of the target prediction block. The block merging described herein is described in Non Patent Literature 1.

[0010] In inter-picture prediction in the standards such as MPEG-1, 2 and MPEG-4, each picture is partitioned into a set of rectangular blocks without an overlap among them and a motion vector is associated with each of the blocks. The motion vector is one obtained by motion search for each block and represents a horizontal displacement and a vertical displacement of a current block from a second block used for prediction of the image signal of the current block.

[0011] Patent Literature 4 below describes a method for carrying out motion compensation prediction with higher accuracy in situations where a boundary of motion exists in an oblique direction in a block. This method is to further partition a block into nonrectangular sub-partitions and perform motion compensated prediction for each of the sub-partitions.

[0012] Patent Literature 5 below describes a method of further partitioning a block into small rectangular sub-partitions and performing motion compensation prediction for each of the sub-partitions. In this method, for encoding a motion vector of a processing target sub-partition, a motion vector predictor is generated from a motion vector of a block being in contact with the processing target sub-partition and located prior in a processing order to the sub-partition, and a difference between the motion vector of the processing target sub-partition and the motion vector predictor, i.e., a differential motion vector only is encoded. In this method, if the processing target sub-partition has no contact with a block prior in the processing order, the motion vector predictor of the processing target sub-partition is generated from a motion vector of another sub-partition prior in the processing order in the block including the processing target sub-partition.

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

Patent Literature 4: Japanese Patent Application Laid-open No. 2005-277968

Patent Literature 5: Japanese Patent Application Laid-open No. 2009-246972

Non Patent Literature

[0014] Non Patent Literature 1: Test Model under Consideration, Joint

Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, 1st Meeting: Dresden, DE, 15-23 April, 2010, Document: JCTVC-A205

Summary of Invention

Technical Problem

[0015] In Non Patent Literature 1 above, candidates for motion information to be used in block merging of a plurality of prediction blocks resulting from partition of a target encoding block serving as an encoding target are selected by the same method, irrespective of the prediction blocks and surrounding situations. For this reason, for example as shown in Fig. (B) of Fig. 2, candidates for motion information in generation of the predicted signal of prediction block T2 include the motion information of prediction block T1 in the same encoding block. The prediction block partitioning type consisting of the prediction block T1 and prediction block T2 is prepared on the assumption that the predicted signals of the two blocks are generated using different pieces of motion information. Therefore, it is unfavorable that the motion information of prediction block T1 is included in the candidates for motion information of prediction block T2. Namely, it can result in inefficient coding.

[0016] It is thus an object of the present invention, in some aspects, to provide an 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 in which candidates for motion information to be used in generation of the predicted signal of the target

prediction block are selected based on previously-encoded or previously-decoded prediction information (motion information and prediction block partitioning type) of the target coding block and surrounding coding blocks, to suppress occurrence of inefficient coding. Namely, in these aspects, the object of the present invention is to 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 achieving improvement in coding efficiency.

[0017] Furthermore, there are the methods of carrying out the motion compensation prediction for each of sub-partitions obtained by partition of a processing target block, as described in Patent Literature 4 or Patent Literature 5. In this motion compensation prediction, it is preferable to generate the motion vector predictor for each sub-partition, based on a motion vector of a block prior in the processing order to a processing target sub-partition, and to encode only the differential motion vector between the motion vector of the sub-partition and the motion vector predictor, in terms of code amount.

[0018] Fig. 23 is a drawing for explaining the motion compensation prediction. As shown in Fig. 23, a processing target block P may have a sub-partition SP1 in contact with at least one block CP prior in a processing order to the block P, and a sub-partition SP2 having no contact with the block CP. A motion vector V2 of such sub-partition SP2 is encoded as it is, without use of the motion vector predictor, in the method described in Patent Literature 4. This method is equivalent to

a method of setting the motion vector predictor to a zero vector.

[0019] On the other hand, in the method described in Patent Literature 5, a motion vector predictor of sub-partition SP2 is generated from the motion vector V1 of the sub-partition SP1 being another sub-partition in the block P and being prior in the processing order to the sub-partition SP2. However, the motion vector of the sub-partition SP1 and the motion vector of the sub-partition SP2 are originally considered to be different from each other. Therefore, the method described in Patent Literature 5 may fail to efficiently encode the motion vector of the sub-partition SP2.

[0020] It is therefore an object of the present invention, in some other aspects, also to provide a video encoding device, video encoding method, and video encoding program capable of achieving improvement in coding efficiency and a video decoding device, video decoding method, and video decoding program corresponding to the video encoding.

Solution to Problem

[0021] The first aspect of the present invention relates to image predictive encoding.

[0022] An image predictive encoding device according to the first aspect of the present invention comprises: region partition means which partitions an input image into a plurality of regions; prediction information estimation means: which sub-partitions a target region serving as an encoding target resulting from partition by the region partition means into a plurality of prediction regions, which determines a prediction block partitioning type indicating a number and region

shapes of the prediction regions suitable for the target region, which predicts motion information for acquisition of each of signals highly correlated with the respective prediction regions from a previously-reconstructed signal, and which obtains prediction information containing the prediction block partitioning type and the motion information; prediction information encoding means which encodes the prediction information associated with the target region; predicted signal generation means which generates a predicted signal of the target region, based on the prediction information associated with the target region; residual signal generation means which generates a residual signal based on the predicted signal of the target region and a pixel signal of the target region; residual signal encoding means which encodes the residual signal generated by the residual signal generation means; residual signal restoration means which decodes encoded data of the residual signal to generate a reconstructed residual signal; and recording means which adds the predicted signal to the reconstructed residual signal to generate a restored pixel signal of the target region, and which stores the restored pixel signal as the previously-reconstructed signal. The prediction information encoding means is configured as follows: the prediction information encoding means has prediction information storage means which stores previously-encoded prediction information; the prediction information encoding means encodes the prediction block partitioning type of the target region and stores the prediction block partitioning type into the prediction information storage means; based on: the prediction information of a neighboring region neighboring the target region, the

number of previously-encoded prediction regions in the target region, and the previously-encoded prediction information of the target region; the prediction information encoding means selects a candidate for motion information to be used in generation of a predicted signal of a target prediction region serving as a next prediction region, from previously-encoded motion information of a region neighboring the target prediction region; according to the number of candidates for the motion information selected, the prediction information encoding means encodes merging block information indicative of generation of the predicted signal of the target prediction region using the selected candidate for motion information and the motion information detected by the prediction information estimation means, or encodes either one of the merging block information and the motion information, and stores the motion information to be used in generation of the predicted signal of the target prediction region, into the prediction information storage means.

[0023] An image predictive encoding method according to the first aspect of the present invention comprises: a region partition step of partitioning an input image into a plurality of regions; a prediction information estimation step of sub-partitioning a target region serving as an encoding target resulting from partition in the region partition step, into a plurality of prediction regions, determining a prediction block partitioning type indicating a number and region shapes of prediction regions suitable for the target region, estimating motion information for acquisition of each of signals highly correlated with the respective prediction regions, from a previously-reconstructed signal, and

obtaining prediction information containing the prediction block partitioning type and the motion information; a prediction information encoding step of encoding the prediction information associated with the target region; a predicted signal generation step of generating a predicted signal of the target region, based on the prediction information associated with the target region; a residual signal generation step of generating a residual signal based on the predicted signal of the target region and a pixel signal of the target region; a residual signal encoding step of encoding the residual signal generated in the residual signal generation step; a residual signal restoration step of decoding encoded data of the residual signal to generate a reconstructed residual signal; and a recording step of adding the predicted signal to the reconstructed residual signal to generate a restored pixel signal of the target region, and storing the restored pixel signal as the previously-reconstructed signal. The prediction information encoding step is configured as follows: the prediction information encoding step comprises encoding the prediction block partitioning type of the target region and storing the prediction block partitioning type into prediction information storage means storing previously-encoded prediction information; the prediction information encoding step comprises, based on the prediction information of a neighboring region neighboring the target region, the number of previously-encoded prediction regions in the target region, and the previously-encoded prediction information of the target region, selecting a candidate for motion information to be used in generation of a predicted signal of a target prediction region serving as a next prediction region, from previously-encoded motion information of a

region neighboring the target prediction region; the prediction information encoding step comprises, according to the number of candidates for motion information selected, encoding merging block information indicative of generation of the predicted signal of the target prediction region using the selected candidate for motion information and the motion information detected in the prediction information estimation step, or encoding either one of the merging block information and the motion information, and storing the motion information to be used in generation of the predicted signal of the target prediction region, into the prediction information storage means.

[0024] An image predictive encoding program according to the first aspect of the present invention causes a computer to function as each of the means of the image predictive encoding device described above.

[0025] According to the first aspect of the present invention, the candidate for motion information to be used in generation of the predicted signal of the target prediction block is selected based on the previously-encoded prediction information (motion information and prediction block partitioning type) of the target encoding block and surrounding encoding blocks, which suppresses occurrence of inefficient coding.

[0026] In an embodiment, based on the number of previously-encoded prediction regions in the target region, the prediction block partitioning type of the target region, and the prediction block partitioning type of the neighboring region neighboring the target region, the candidate for motion information of the target prediction region serving as the next prediction region may be selected from the previously-encoded motion

information of the region neighboring the target prediction region.

[0027] In an embodiment, based on the number of previously-encoded prediction regions in the target region and the prediction block partitioning type of the target region, the candidate for motion information of the target prediction region serving as the next prediction region may be selected from the previously-encoded motion information of the region neighboring the target prediction region; when the target region is partitioned into two prediction regions and when the target prediction region is a prediction region to be encoded second in the target region, motion information of a region which is neighboring the target prediction region and which is not included in the target region may be selected as the candidate for motion information to be used in generation of the predicted signal of the target prediction region.

[0028] In an embodiment, based on: the number of previously-encoded prediction regions in the target region, the prediction block partitioning type of the target region, the previously-encoded motion information in the target region, and the motion information of the neighboring region neighboring the target region; the candidate for motion information to be used in generation of the predicted signal of the target prediction region serving as the next prediction region may be selected from the previously-encoded motion information of the region neighboring the target prediction region; when the target region is partitioned into two prediction regions, when the target prediction region is a prediction region to be encoded second in the target region, and when the motion information of the prediction region encoded first in the target region is the same as motion information of a region which is neighboring the

target prediction region and which is not included in the target region, it may be determined that the motion information of the region neighboring the target prediction region is not used in generation of the predicted signal of the target prediction region, and the motion information may be encoded.

[0029] The second aspect of the present invention relates to image predictive decoding.

[0030] An image predictive decoding device according to the second aspect of the present invention comprises: data analysis means which extracts, from compressed data of an image resulting from partition into a plurality of regions, and encoding: encoded data of prediction information to indicate a prediction method to be used in prediction of a signal of a target region serving as a decoding target, encoded data of a predicted signal of the target region, and encoded data of a residual signal; prediction information decoding means which decodes the encoded data of the prediction information to restore a prediction block partitioning type indicating a number and region shapes of prediction regions that are sub-partitioned regions of the target region, and motion information for acquisition of each of predicted signals of the respective prediction regions from a previously-reconstructed signal; predicted signal generation means which generates the predicted signal of the target region, based on the prediction information associated with the target region; residual signal restoration means which restores a reconstructed residual signal of the target region from the encoded data of the residual signal; and recording means which adds the predicted signal to the reconstructed residual signal to restore a pixel signal of the

target region, and which stores the pixel signal as the previously-reconstructed signal. The prediction information decoding means is configured as follows: the prediction information decoding means has prediction information storage means which stores previously-decoded prediction information; the prediction information decoding means decodes the prediction block partitioning type of the target region and stores the prediction block partitioning type into the prediction information storage means; based on the prediction information of a neighboring region neighboring the target region, the number of previously-decoded prediction regions in the target region, and the previously-decoded prediction information of the target region, the prediction information decoding means selects a candidate for motion information to be used in generation of a predicted signal of a target prediction region as a next prediction region, from previously-decoded motion information of a region neighboring the target prediction region; according to the number of candidates for motion information selected, the prediction information decoding means decodes merging block information indicative of generation of the predicted signal of the target prediction region using the selected candidate for motion information and the motion information, or either decodes one of the merging block information and the motion information, and stores the motion information to be used in generation of the predicted signal of the target prediction region, into the prediction information storage means.

[0031] An image predictive decoding method according to the second aspect of the present invention comprises: a data analysis step of

extracting, from compressed data of an image resulting from partition into a plurality of regions and encoding: encoded data of prediction information to indicate a prediction method to be used in prediction of a signal of a target region serving as a decoding target, encoded data of a predicted signal of the target region, and encoded data of a residual signal; a prediction information decoding step of decoding the encoded data of the prediction information to restore a prediction block partitioning type indicating a number and region shapes of prediction regions that are sub-partitioned regions of the target region, and motion information for acquisition of each of predicted signals of the respective prediction regions from a previously-reconstructed signal; a predicted signal generation step of generating the predicted signal of the target region, based on the prediction information associated with the target region; a residual signal restoration step of restoring a reconstructed residual signal of the target region from the encoded data of the residual signal; and a recording step of adding the predicted signal to the reconstructed residual signal to restore a pixel signal of the target region, and storing the pixel signal as the previously-reconstructed signal. The prediction information decoding step is configured as follows: the prediction information decoding step comprises decoding the prediction block partitioning type of the target region and storing the prediction block partitioning type as previously-decoded prediction information into prediction information storage means that stores previously-decoded prediction information; the prediction information decoding step comprises, based on: the prediction information of a neighboring region neighboring the target region, the number of

previously-decoded prediction regions in the target region, and the previously-decoded prediction information of the target region; selecting a candidate for motion information to be used in generation of a predicted signal of a target prediction region serving as a next prediction region, from previously-decoded motion information of a region neighboring the target prediction region; and the prediction information decoding step comprises, according to the number of candidates for motion information selected, decoding merging block information indicative of generation of the predicted signal of the target prediction region using the selected candidate for motion information and the motion information, or decoding either one of the merging block information and the motion information, and storing the motion information to be used in generation of the predicted signal of the target prediction region into the prediction information storage means.

[0032] An image predictive decoding program according to the second aspect of the present invention causes a computer to function as each of the means of the image predictive decoding device described above.

[0033] According to the second aspect of the present invention, an image can be decoded from the compressed data generated by the aforementioned image predictive encoding.

[0034] In an embodiment, based on the number of previously-decoded prediction regions in the target region, the prediction block partitioning type of the target region, and the prediction block partitioning type of the neighboring region neighboring the target region, the candidate for motion information of the target prediction region serving as the next prediction region may be selected, from the previously-decoded motion

information of the region neighboring the target prediction region.

[0035] In an embodiment, based on the number of previously-decoded prediction regions in the target region and the prediction block partitioning type of the target region, the candidate for motion information to be used in generation of the predicted signal of the target prediction region serving as the next prediction region may be selected from the previously-decoded motion information of the region neighboring the target prediction region; when the target region is partitioned into two prediction regions and when the target prediction region is a prediction region to be decoded second in the target region, motion information of a region which is neighboring the target prediction region and which is not included in the target region may be selected as the candidate for motion information of the target prediction region.

[0036] In an embodiment, based on the number of previously-decoded prediction regions in the target region, the prediction block partitioning type of the target region, the previously-decoded motion region in the target region, and the motion information of the neighboring region neighboring the target region, the candidate for motion information to be used in generation of the predicted signal of the target prediction region serving as the next prediction region may be selected, from the previously-decoded motion information of the region neighboring the target prediction region; when the target region is partitioned into two prediction regions, when the target prediction region is a prediction region to be decoded second in the target region, and when the motion information of the prediction region decoded first in the target region is

the same as motion information of a region which is neighboring the target prediction region and which is not included in the target region; it may be determined that the motion information of the region neighboring the target prediction region is not used in generation of the predicted signal of the target prediction region, and the motion information may be decoded.

[0037] The third aspect of the present invention relates to video encoding.

[0038] A video encoding device according to the third aspect comprises partition means, sub-partition generation means, motion detection means, predicted signal generation means, motion prediction means, differential motion vector generation means, residual signal generation means, addition means, storage means, and encoding means. The partition means partitions an input image in a video sequence into a plurality of partitions. The sub-partition generation means partitions a processing target partition generated by the partition means, into a plurality of sub-partitions and generates shape information to specify shapes of the sub-partitions. The motion detection means detects a motion vector of the processing target partition. The predicted signal generation means generates a predicted signal of the processing target partition from a previously-reconstructed image signal, using the motion vector detected by the motion detection means. The motion prediction means generates a motion vector predictor of the processing target partition, based on the shape information generated by the sub-partition generation means and a motion vector of a previously-processed partial region. The previously-processed partial region can be a partition or a

sub-partition prior in a processing order to the processing target partition. The differential motion vector generation means generates a differential motion vector, based on a difference between the motion vector used in generation of the predicted signal of the processing target partition and the motion vector predictor. The residual signal generation means generates a residual signal based on a difference between the predicted signal and a pixel signal of the processing target partition. The addition means adds the residual signal to the predicted signal to generate a reconstructed image signal. The storage means stores the reconstructed image signal as a previously-reconstructed image signal. The encoding means encodes: the residual signal generated by the residual signal generation means, the differential motion vector generated by the differential vector generation means, and the shape information generated by the sub-partition generation means; to generate compressed data. When a processing target sub-partition in the processing target partition has no contact with a partition prior in the processing order to the processing target sub-partition, the motion prediction means generates a motion vector predictor of the processing target sub-partition, based on a motion vector of a previously-processed partial region belonging to a domain that also contains the processing target sub-partition or an other domain. The domain and the other domain can be partitioned by an extension line of a boundary between the processing target sub-partition and another sub-partition in the processing target partition.

[0039] A video encoding method according to the third aspect comprises: (a) a partition step of partitioning an input image from a

video sequence into a plurality of partitions; (b) a sub-partition generation step of partitioning a processing target partition generated in the partition step, into a plurality of sub-partitions and generating shape information to specify shapes of the sub-partitions; (c) a motion detection step of detecting a motion vector of the processing target partition; (d) a predicted signal generation step of generating a predicted signal of the processing target partition from a previously-reconstructed image signal, using the motion vector detected in the motion detection step; (e) a motion prediction step of generating a motion vector predictor of the processing target partition, based on the shape information generated in the sub-partition generation step and a motion vector of a previously-processed partial region as a partition or a sub-partition prior in a processing order to the processing target partition; (f) a differential motion vector generation step of generating a differential motion vector, based on a difference between the motion vector used in generation of the predicted signal of the processing target partition and the motion vector predictor; (g) a residual signal generation step of generating a residual signal based on a difference between the predicted signal and a pixel signal of the processing target partition; (h) an addition step of adding the residual signal to the predicted signal to generate a reconstructed image signal; (i) a storage step of storing the reconstructed image signal as a previously-reconstructed image signal; and (j) an encoding step of encoding: the residual signal generated in the residual signal generation step, the differential motion vector generated in the differential motion vector generation step, and the shape information generated in the

sub-partition generation step; to generate compressed data. When a processing target sub-partition in the processing target partition has no contact with a partition prior in the processing order to the processing target sub-partition, the motion prediction step comprises generating a motion vector predictor of the processing target sub-partition, based on a motion vector of a previously-processed partial region which belongs to either a domain containing the processing target sub-partition or an other domain. The domain and the other domain may be partitioned by an extension line of a boundary between the processing target sub-partition and another sub-partition in the processing target partition.

[0040] A video encoding program according to the third aspect causes a computer to function as each of the means of the video encoding device described above.

[0041] The domain including a sub-partition having no contact with a partition prior in the processing order, out of the two domains defined by the aforementioned extension line of the boundary, is highly likely to have a motion similar to a motion of the sub-partition. Therefore, according to the third aspect, the accuracy of the motion vector predictor improves, the value of the differential motion vector becomes smaller, and the motion vector is encoded with a smaller code amount. Therefore, the encoding efficiency improves.

[0042] The fourth aspect of the present invention relates to video decoding.

[0043] A video decoding device according to the fourth aspect comprises decoding means, motion prediction means, vector addition means, predicted signal generation means, addition means, and storage

means. The decoding means decodes compressed data to generate a reconstructed residual signal of a processing target partition in an image, a differential motion vector of the processing target partition, and shape information to specify shapes of a plurality of sub-partitions in the processing target partition. The motion prediction means generates a motion vector predictor of the processing target partition, based on the shape information and a motion vector of a previously-processed partial region that is a partition or a sub-partition prior in a processing order to the processing target partition. The vector addition means adds the motion vector predictor generated by the motion prediction means, to the differential motion vector generated by the decoding means, to generate a motion vector of the processing target partition. The predicted signal generation means generates a predicted signal of the processing target partition from a previously-reconstructed image signal, based on the motion vector of the processing target partition. The addition means adds the predicted signal to the reconstructed residual signal generated by the decoding means, to generate a reconstructed image signal. The storage means stores the reconstructed image signal as a previously-reconstructed image signal. When a processing target sub-partition in the processing target partition has no contact with a partition prior in the processing order to the processing target sub-partition, the motion prediction means generates a motion vector predictor of the processing target sub-partition, based on a motion vector of a previously-processed partial region belonging to either a domain containing the processing target sub-partition or an other domain. The domain and the other domain can be partitioned by an

extension line of a boundary between the processing target sub-partition and another sub-partition in the processing target partition.

[0044] A video decoding method according to the fourth aspect is a method for decoding compressed data to generate a video sequence, comprising: (a) a decoding step of decoding the compressed data to generate a reconstructed residual signal of a processing target partition in an image, a differential motion vector of the processing target partition, and shape information to specify shapes of a plurality of sub-partitions in the processing target partition; (b) a motion prediction step of generating a motion vector predictor of the processing target partition, based on the shape information and a motion vector of a previously-processed partial region as a partition or a sub-partition prior in a processing order to the processing target partition; (c) a vector addition step of adding the motion vector predictor generated in the motion prediction step, to the differential motion vector generated in the decoding step, to generate a motion vector of the processing target partition; (d) a predicted signal generation step of generating a predicted signal of the processing target partition from a previously-reconstructed image signal, based on the motion vector of the processing target partition; (e) an addition step of adding the predicted signal to the reconstructed residual signal generated in the decoding step, to generate a reconstructed image signal; and (f) a storage step of storing the reconstructed image signal as a previously-reconstructed image signal. When a processing target sub-partition in the processing target partition has no contact with a partition prior in the processing order to the processing target sub-partition, the motion prediction step comprises

generating a motion vector predictor of the processing target sub-partition, based on a motion vector of a previously-processed partial region belonging to either a domain containing the processing target sub-partition or an other domain. The domain and the other domain being partitioned by an extension line of a boundary between the processing target sub-partition and another sub-partition in the processing target partition.

[0045] A video decoding program according to the fourth aspect causes a computer to function as each of the means of the video decoding device described above.

[0046] According to the fourth aspect, the motion vector predictor of the sub-partition is generated from a previously-decoded motion vector in the domain containing the sub-partition which has no contact with a partition prior in the processing order. This motion vector predictor is highly likely to be similar to the motion vector of the sub-partition. According to the forgoing embodiments, therefore, the accuracy of the motion vector predictor improves, the value of the differential motion vector becomes smaller, and it becomes feasible to perform decoding from compressed data with a smaller bit amount. Therefore, efficient decoding is achieved.

Advantageous Effects of Invention

[0047] 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 according to some aspects of the present invention provide the effect of more efficient encoding of motion

information because the candidate for motion information to be used in generation of the predicted signal of the target prediction block can be selected based on the surrounding previously-encoded or previously-decoded information.

[0048] Some other aspects of the present invention provide the video encoding device, video encoding method, and video encoding program capable of improving coding efficiency. Furthermore, there are provided the video decoding device, video decoding method, and video decoding program corresponding to the foregoing video encoding.

Brief Description of Drawings

[0049] Fig. 1 is a block diagram showing an image predictive encoding device according to an embodiment.

Fig. 2 is a schematic diagram for explaining candidates for motion information in the conventional block merging.

Fig. 3 is a schematic diagram for explaining candidates for motion information in block merging according to an embodiment.

Fig. 4 is a flowchart to illustrate a processing procedure of a prediction information encoder shown in Fig. 1.

Fig. 5 is a flowchart showing a procedure of an image predictive encoding method of the image predictive encoding device shown in Fig. 1.

Fig. 6 is a block diagram showing an image predictive decoding device according to an embodiment.

Fig. 7 is a flowchart to illustrate a processing procedure of a prediction information decoder shown in Fig. 6.

Fig. 8 is a flowchart showing a procedure of an image predictive

decoding method of the image predictive decoding device shown in Fig. 6.

Fig. 9 is a first schematic drawing for explaining processes using motion information of a plurality of neighboring blocks neighboring a target prediction block, as motion information of the target prediction block.

Fig. 10 is a second schematic drawing for explaining processes using motion information of a plurality of neighboring blocks neighboring a target prediction block, as motion information of the target prediction block.

Fig. 11 is a flowchart to illustrate a process using motion information of a plurality of neighboring blocks neighboring a target prediction block, as motion information of the target prediction block.

Fig. 12 is a third schematic drawing for explaining processes using motion information of a plurality of neighboring blocks neighboring a target prediction block, as motion information of the target prediction block.

Fig. 13 is a second example of a flowchart to illustrate a process using motion information of a plurality of neighboring blocks neighboring a target prediction block, as motion information of the target prediction block.

Fig. 14 is a second example of a schematic diagram for explaining candidates for motion information in block merging according to an embodiment.

Fig. 15 is a third example of a schematic diagram for explaining candidates for motion information in block merging according to an

embodiment.

Fig. 16 is a block diagram showing a program capable of executing the image predictive encoding method according to an embodiment.

Fig. 17 is a block diagram showing a program capable of executing the image predictive decoding method according to an embodiment.

Fig. 18 is a drawing showing a hardware configuration of a computer for executing a program recorded in a recording medium.

Fig. 19 is a perspective view of a computer for executing a program stored in a recording medium.

Fig. 20 is a schematic diagram for explaining prediction block partitioning types of encoding blocks.

Fig. 21 is a schematic diagram about a motion estimation process (A) and a template matching process (B) in inter-picture prediction.

Fig. 22 is a schematic diagram for explaining the conventional intra-picture prediction method.

Fig. 23 is a drawing for explaining the motion compensated prediction.

Fig. 24 is a drawing schematically showing a configuration of a video encoding device according to an embodiment.

Fig. 25 is a drawing for explaining generation of sub-partitions.

Fig. 26 is a drawing showing a configuration of a motion predictor in an embodiment.

Fig. 27 is a flowchart of a video encoding method according to

an embodiment.

Fig. 28 is a flowchart showing a process of the motion predictor according to an embodiment.

Fig. 29 is a drawing showing an example of sub-partitions of a target partition and surrounding partial regions.

Fig. 30 is a drawing showing another example of sub-partitions of a target block and surrounding partial regions.

Fig. 31 is a drawing showing still other examples of sub-partitions of a target block and surrounding partial regions.

Fig. 32 is a drawing showing still another example of sub-partitions of a target block and a surrounding partial region.

Fig. 33 is a drawing showing still another example of sub-partitions of a target block and surrounding partial regions.

Fig. 34 is a drawing schematically showing a configuration of a video decoding device according to an embodiment.

Fig. 35 is a drawing showing a configuration of a motion predictor according to an embodiment.

Fig. 36 is a flowchart of a video decoding method according to an embodiment.

Fig. 37 is a flowchart showing a process of the motion predictor according to an embodiment.

Fig. 38 is a drawing showing a configuration of a video encoding program according to an embodiment.

Fig. 39 is a drawing showing a configuration of a video decoding program according to an embodiment.

Description of Embodiments

[0050] A variety of embodiments will be described below in detail with reference to the accompanying drawings. In the description of the drawings identical or equivalent elements will be denoted by the same reference signs, without redundant description.

[0051] Fig. 1 is a block diagram showing an image predictive encoding device 100 according to an embodiment. This image predictive encoding device 100 is provided with input terminal 101, block partitioner 102, predicted signal generator 103, frame memory 104, subtracter 105, transformer 106, quantizer 107, inverse quantizer 108, inverse transformer 109, adder 110, quantized transform coefficient encoder 111, output terminal 112, prediction block partitioning type selector 113, motion information estimator 114, prediction information memory 115, and prediction information encoder 116.

[0052] The transformer 106, quantizer 107, and quantized transform coefficient encoder 111 function as residual signal encoding means and the inverse quantizer and inverse transformer function as residual signal restoration means. The prediction block partitioning type selector 113 and motion information estimator 114 function as prediction information estimation means, and the prediction information memory 115 and prediction information encoder 116 as prediction information encoding means.

[0053] The input terminal 101 is a terminal that accepts input of a signal of a video sequence consisting of a plurality of images.

[0054] The block partitioner 102 partitions an image serving as an encoding target represented by a signal input from the input terminal 101, into a plurality of regions (encoding blocks). In the present

embodiment the encoding target image is partitioned into blocks each consisting of 16×16 pixels, but the image may be partitioned into blocks of any other size or shape. Furthermore, blocks of different sizes may be mixed in a frame.

[0055] The prediction block partitioning type selector 113 partitions a target region (target encoding block) serving as an encoding target into prediction regions to be subjected to a prediction process. For example, it selects one of (A) to (H) in Fig. 20 for each encoding block and sub-partitions the encoding block according to the selected mode. Each partitioned region is called a prediction region (prediction block) and each of the partition methods (A) to (H) in Fig. 20 is called a prediction block partition type. An available method of selecting a prediction block partitioning type is, for example, a method of carrying out each of sub-partitions of the signal of the target encoding block fed via line L102, actually carrying out below-described prediction processing and encoding processing, and selecting a partitioning type to minimize a rate-distortion value calculated from the power of an encoding error signal between the original signal of the encoding block and a reconstructed signal, and a code amount necessary for encoding of the encoding block, but is not limited thereto. The prediction block partitioning type of the target encoding block is output via line L113a, via line L113b, and via line L113c to the prediction information memory 115, to the motion information estimator 114, and to the predicted signal generator 103, respectively.

[0056] The motion information estimator 114 detects motion information necessary for generation of a predicted signal of each

prediction block in the target encoding block. An applicable method of generation of the predicted signal (prediction method) are inter-picture prediction and the intra-picture prediction (the intra-picture prediction is not shown) as described in the background, but are not limited thereto. It is assumed herein that the motion information is detected by the block matching shown in Fig. 21. An original signal of the target prediction block being a prediction target can be generated from the original signal of the encoding block fed via line L102a and the prediction block partitioning type of the target encoding block fed via line L113b. A predicted signal to minimize the sum of absolute differences for the original signal of the target prediction block is detected from image signals acquired via line L104. In this case, the motion information contains a motion vector, an inter-picture prediction mode (forward/backward/bidirectional prediction), a reference frame number, and so on. The detected motion information is output each via line L114 to the prediction information memory 115 and the prediction information encoder 116.

[0057] The prediction information memory 115 stores the input motion information and prediction block partitioning type.

[0058] The prediction information encoder 116 selects candidates for motion information to be used in block merging of each prediction block, entropy-encodes prediction information of the target encoding block, and outputs the encoded data via line L116 to the output terminal 112. An applicable method of entropy encoding includes the arithmetic coding, the variable length coding, and so on, but is not limited thereto. The prediction information contains block merging

information for execution of block merging using motion information of blocks neighboring the prediction block, in addition to the prediction block partitioning type of the target encoding block and the motion information of the prediction block. The process of the prediction information encoder 116 will be described later.

[0059] The predicted signal generator 103 acquires previously-reconstructed signals from the frame memory 104, based on the motion information of each prediction block in the target encoding block fed via line L114 and the prediction block partitioning type fed via line L113c, and generates a predicted signal of each prediction block in the target encoding block.

[0060] The predicted signal generated in the predicted signal generator 103 is output via line L103 to the subtracter 105 and the adder 110.

[0061] The subtracter 105 subtracts the predicted signal for the target encoding block fed via line L103, from the pixel signal of the target encoding block fed via line L102b after the partition by the block partitioner 102, to generate a residual signal. The subtracter 105 outputs the residual signal obtained by the subtraction, via line L105 to the transformer 106.

[0062] The transformer 106 is a part that performs a discrete cosine transform on the input residual signal. The quantizer 107 is a part that quantizes transform coefficients obtained by the discrete cosine transform by the transformer 106. The quantized transform coefficient encoder 111 entropy-encodes the quantized transform coefficients obtained by the quantizer 107. The encoded data is output via line L111 to the output terminal 112. An applicable method of entropy

encoding includes the arithmetic coding, the variable length coding, and so on, but is not limited thereto.

[0063] The output terminal 112 outputs the pieces of information input from the prediction information encoder 116 and the quantized transform coefficient encoder 111, together to the outside.

[0064] The inverse quantizer 108 performs inverse quantization on the quantized transform coefficients. The inverse transformer 109 performs an inverse discrete cosine transform to restore a residual signal. The adder 110 adds the restored residual signal to the predicted signal fed via line L103, to reconstruct a signal of the target encoding block, and stores the reconstructed signal into the frame memory 104. The present embodiment employs the transformer 106 and the inverse transformer 109, but another transform process may be employed instead of these transformers. Furthermore, the transformer 106 and the inverse transformer 109 are not always essential. In this manner, for use in generation of the predicted signal of the subsequent target encoding block, the reconstructed signal of the target encoding block subjected to encoding is restored by the inverse process to be stored in the frame memory 104.

[0065] Next, the process of the prediction information encoder 116 will be described. The prediction information encoder 116 first selects candidates for motion information to be used in block merging of each prediction block (candidates for motion information to be used in generation of a predicted signal of a target prediction region) from the motion information of blocks neighboring the target prediction block. The block merging refers to generation of the predicted signal of the

target prediction block using the motion information of neighboring blocks. Next, the prediction information encoder 116 compares the motion information detected by the motion information estimator 114, with the candidates for motion information thus selected, to determine whether block merging is to be carried out. Then, according to the number of candidates for motion information to be used in block merging, and the applicability of execution of block merging, the prediction information encoder 116 entropy-encodes either or both of block merging information and the motion information, along with the prediction block partitioning type. The block merging information includes merging identification information (merge_flag) to indicate whether the predicted signal of the target prediction block is to be generated using motion information of a neighboring block, i.e., whether block merging is to be carried out, and merging block selection information (merge_flag_left) to indicate which should be used between pieces of motion information of two or more blocks neighboring the target prediction block to generate the predicted signal of the target prediction block.

[0066] If there is no candidate for motion information to be used in block merging of each prediction block, there is no need for encoding these two pieces of information, i.e., the merging identification information and the merging block selection information. If there is one candidate for motion information, the merging identification information is encoded; if there are two or more candidates for motion information and if block merging is carried out, the two pieces of information, i.e., the pieces of merging identification information and

merging block selection information are encoded. Even with two or more candidates for motion information, there is no need for encoding of the merging block selection information if block merging is not carried out.

[0067] Fig. 3 is a schematic diagram for explaining a process of selecting candidates for motion information to be used in block merging of a prediction block according to an embodiment. Fig. 3 shows an example of a prediction block partitioning type to vertically bisect the encoding block (or to partition the encoding block into right and left blocks), as in the case of the block 301 shown in (B) of Fig. 20. The below will describe this block 301 as an example, but the same description also applies to blocks 302, 304, 305, 306, and 307.

[0068] Selection of a candidate for motion information is carried out based on information below.

1) The number of already-encoded/already-decoded prediction blocks in the target encoding block

2) The prediction block partitioning type of the target encoding block

3) The prediction block partitioning type of a block neighboring the target prediction block

4) The motion information of the already-encoded/already-decoded prediction blocks in the target encoding block

5) The motion information and the prediction mode (intra-picture prediction/inter-picture prediction) of the block neighboring the target prediction block

[0069] In the example of Fig. 3, a candidate for motion information to be used in block merging is selected using the pieces of information of

1), 2), 4), and 5).

[0070] First, it is found from the information of 2) that the total number of prediction blocks in the target encoding block 400 is 2, prediction blocks T1 and T2, and that the encoding block is vertically partitioned into two sub-partitions. It is found from the information of 1) that the next prediction block is either prediction block T1 or prediction block T2.

[0071] When the next prediction block is the prediction block T1 (or when the number of already-encoded/already-decoded prediction blocks in the target encoding block is 0), pieces of motion information of neighboring block A and neighboring block B are set as candidates for motion information for block merging (arrows in the drawing indicate that pieces of motion information of neighboring blocks A and B are candidates for motion information to be used in generation of the predicted signal of prediction block T1). On this occasion, if the neighboring block A or B is a block generated by intra-picture prediction or a block outside the picture, the motion information of the block may be excluded from candidates for motion information for block merging (it is also possible to set the motion information to a pseudo default value; for example, the motion vector is set to 0 and the reference frame number to 0). If pieces of motion information of the two neighboring blocks A and B are identical with each other, the motion information of one neighboring block may be excluded from the candidates.

[0072] When the next prediction block is the prediction block T2 (or when the number of already-encoded/already-decoded prediction blocks

in the target encoding block is 1), as shown in (A) of Fig. 3, the motion information of the neighboring block T1 is excluded from candidates for motion information for block merging. This is because the target encoding block is originally partitioned into two blocks on the assumption that the predicted signals of the prediction block T1 and the prediction block T2 are generated from different pieces of motion information. Namely, this is for avoiding such a situation that the motion information of the prediction block T2 becomes the same as the motion information of the prediction block T1. Since this process provides only one piece of motion information for block merging of the prediction block T2, the cost for encoding of merging block selection information can be reduced (an arrow in the drawing indicates that the motion information of neighboring block D is applicable to generation of the predicted signal of the prediction block T2).

[0073] Furthermore, based on the pieces of information of 4) and 5) above, the motion information of the prediction block T1 is compared with the motion information of the neighboring block D and if these pieces of motion information of the prediction block T1 and neighboring block D are identical with each other, the motion information of the neighboring block D is also excluded from candidates for motion information for block merging as shown in (B) of Fig. 3. The reason for it is that if the predicted signal of the prediction block T2 is generated using the motion information of the neighboring block D, pieces of motion information of the prediction blocks T1 and T2 become identical with each other. Owing to this process, the motion information for block merging of the prediction block T2

becomes null, which can reduce the cost for encoding of merging identification information and merging block selection information.

[0074] Fig. 4 is a flowchart of the prediction information encoder 116 to substantialize the process of Fig. 3.

[0075] First, the prediction information encoder 116 encodes the prediction block partitioning type of the target encoding block and stores the prediction block partitioning type into the prediction information memory 115. At the same time, the prediction information encoder 116 sets the number N of prediction blocks in the target encoding block, based on the encoded prediction block partitioning type, and resets a target prediction block number i to 0 (step S151). Next, the prediction information encoder 116 determines whether a target prediction block is a prediction block to be encoded last in the target encoding block and whether the number of prediction blocks in the target encoding block is not less than 2 (step S152). For example, in the case of N=2, the decision is yes with i=1, and the processing proceeds to step S157. In the case of N=4 ((D) of Fig. 20), the decision becomes yes with i=3. When the decision is no, the processing proceeds to step S153. In the case of Fig. 3, the processing proceeds to step S153 when the target prediction block is the prediction block T1; the processing proceeds to step S157 when the target prediction block is the prediction block T2.

[0076] In step S153, the merging identification information is encoded. The merging identification information becomes yes (merge_flag=1, generation of the predicted signal of the target prediction block using a candidate for motion information) if the motion information of the

target prediction block coincides with a candidate for motion information for block merging; otherwise, the merging identification information becomes no (merge_flag=0, generation of the predicted signal of the target prediction block using the encoded motion information). Next, when the motion information of the target prediction block coincides with a candidate for motion information for block merging, the processing proceeds to step S164. In step S164, the prediction information encoder 116 determines whether there are two candidates for motion information, and if the number of candidates for motion information is 2, it encodes the merging block selection information and the processing proceeds to step S155. When the number of candidates for motion information is 1, the processing proceeds to step S165. On the other hand, if the motion information of the target prediction block does not coincide with any candidate for motion information for block merging, the processing proceeds to step S156 and the prediction information encoder 116 encodes the motion information detected by the motion information estimator 114, and then proceeds to step S165.

[0077] In step S157, the prediction information encoder 116 determines whether every piece of already-encoded motion information of the target encoding block coincides with the motion information of a neighboring block that does not belong to the target encoding block. The description of this step S157 means that, in the case of N=2, the motion information of the prediction block T1 shown in Fig. 3 is compared with the motion information of the neighboring block D. Furthermore, the description of step S157 means that, in the case of N=4

((D) of Fig. 20), the target prediction block is the lower right partitioned block and the pieces of motion information of the other three prediction blocks (upper left, upper right, and lower left) are compared with each other. When the determination is yes (or when pieces of motion information in comparison are coincident), the number of candidates for motion information to be used in block merging of the target prediction block is 0, as shown in the example of (B) of Fig. 3, and thus the prediction information encoder 116 encodes the motion information detected by the motion information estimator 114, without transmitting the block merging information, and then the processing proceeds to step S165 (step S160). On the other hand, when the determination is no (or when pieces of motion information in comparison are not coincident), the processing proceeds to step S163. In the case of N=4, pieces of motion information of the upper right and lower left blocks in the target encoding block are those of neighboring blocks to the target prediction block. For this reason, application of block merging to the target prediction block (lower right) with coincidence of the motion information of the three prediction blocks (upper left, upper right, and lower left) means that the predicted signals of the four prediction blocks in the target encoding block are all generated using the same motion information. For this reason, in the case where N=4 and where the pieces of motion information of the three prediction blocks (upper left, upper right, and lower left) are identical with each other, the number of candidates for motion information of the target prediction block (lower right) is set to 0.

[0078] In step S163, the prediction information encoder 116 determines

whether the prediction block partitioning type of the target encoding block is a bisectional type, and if the determination is no, the processing proceeds to step S153 (the description hereinafter is omitted). When the determination in step S163 is yes, the processing proceeds to step S158, in which the prediction information encoder 116 encodes the merging identification information. In this case, since the number of candidates for motion information to be used in block merging of the target prediction block is 1, as in the example of (A) of Fig. 3, there is no need for encoding of the merging block selection information. Next, if the motion information of the target prediction block coincides with the candidate for motion information for block merging, the processing proceeds to step S165. When the motion information of the target prediction block does not coincide with the candidate for motion information for block merging, the processing proceeds to step S160, in which the prediction information encoder 116 encodes the motion information detected by the motion information estimator 114, and then the processing proceeds to step S165.

[0079] In step S165, the motion information of the target block is stored into the prediction information memory 115. Subsequently, in step S161 the prediction information encoder 116 determines whether encoding is completed for all the prediction blocks in the target encoding block (whether i=N−1); when i=N−1, this prediction information encoding processing of the target encoding block is terminated; when $i < N-1$, the number i is updated in step S162 (i=i+1), and the processing returns to step S152 to perform the motion information encoding processing of the next prediction block.

[0080] Since the candidates for motion information to be used in block merging of the prediction block can be selected in advance using the pieces of information below, as described above, it becomes feasible to efficiently transmit the block merging information.

1) The number of already-encoded/already-decoded prediction blocks in the target encoding block

2) The prediction block partitioning type of the target encoding block

4) The motion information of the already-encoded/already-decoded prediction blocks in the target encoding block

5) The motion information and prediction mode (intra-picture prediction/inter-picture prediction) of the block neighboring the target prediction block

[0081] Fig. 5 is a flowchart showing a procedure of an image predictive encoding method in the image predictive encoding device 100 according to the present embodiment. First, the block partitioner 102 partitions an input picture into 16×16 encoding blocks (the picture may be partitioned into blocks of other size or shape, or blocks of different sizes may be mixed in the frame). Then the prediction block partitioning type selector 113 and the motion information estimator 114 determine the prediction block partitioning type of the target encoding block serving as an encoding target and the motion information of each of prediction blocks thereof (step S101). Next, the prediction information encoder 116 encodes the prediction information (step S102, Fig. 4).

[0082] Next, the predicted signal generator 103 generates the predicted signal of the target encoding block, based on the prediction block

partitioning type of the target encoding block and the motion information of each of the prediction blocks, and a residual signal indicative of a difference between a pixel signal of the target encoding block and the predicted signal is transformed and encoded by the transformer 106, quantizer 107, and quantized transform coefficient encoder 111 (step S103). The encoded data of the prediction information and the quantized transform coefficients is then output through the output terminal 112 (step S104).

[0083] For predictive encoding of the subsequent target encoding block, the encoded residual signal is decoded by the inverse quantizer 108 and the inverse transformer 109 after these processes or in parallel with these processes. Then the adder 110 adds the decoded residual signal to the predicted signal to reconstruct a signal of the target encoding block. The reconstructed signal is stored as a reference picture into the frame memory 104 (step S105). If the processing is not completed for all the target encoding blocks, the processing returns to step S101 to carry out the processing for the next target encoding block. When the processing is completed for all the target encoding blocks, the processing is terminated (step S106).

[0084] Next, image predictive decoding according to an embodiment will be described. Fig. 6 is a block diagram showing an image predictive decoding device 200 according to an embodiment. This image predictive decoding device 200 is provided with input terminal 201, data analyzer 202, inverse quantizer 203, inverse transformer 204, adder 205, output terminal 206, quantized transform coefficient decoder 207, prediction information decoder 208, frame memory 104, predicted

signal generator 103, and prediction information memory 115.

[0085] The inverse quantizer 203, inverse transformer 204, and quantized transform coefficient decoder 207 function as residual signal decoding means. The decoding means by the inverse quantizer 203 and the inverse transformer 204 may be implemented using any means other than these. Furthermore, the inverse transformer 204 may be excluded. The prediction information memory 115 and the prediction information decoder 208 function as prediction information decoding means.

[0086] The input terminal 201 accepts input of compressed data resulting from the compression encoding by the aforementioned image predictive encoding method. This compressed data contains the information of quantized transform coefficients resulting from transformation, quantization, and entropy encoding of the residual signal, for each of a plurality of partitioned encoding blocks, and encoded data of the prediction information for generation of the predicted signal of the blocks. The prediction information herein contains the block merging information for execution of block merging using the motion information as candidates for block merging, in addition to the prediction block partitioning type of the target encoded block and the motion information of prediction blocks. Furthermore, the motion information contains the motion vector, the inter-picture prediction mode (forward/backward/bidirectional prediction), the reference frame number, and so on.

[0087] The data analyzer 202 analyzes the compressed data input through the input terminal 201, separates the data about the target

encoded block serving as a decoding target into encoded data of quantized transform coefficients and encoded data of prediction information, and outputs them via line L202a and via line L202b to the quantized transform coefficient decoder 207 and to the prediction information decoder 208, respectively.

[0088] The prediction information decoder 208 selects a candidate for motion information to be used in block merging of each prediction block and entropy-decodes the encoded data of the prediction information associated with the target encoded block. The decoded prediction information is output via line L208a and via line L208b to the predicted signal generator 103 and to the prediction information memory 115, respectively. The prediction information memory 115 stores the input prediction information. The processing of the prediction information decoder 208 will be described later.

[0089] The predicted signal generator 103 acquires previously-reconstructed signals from the frame memory 104, based on the prediction information of the target encoded block fed via line L208a, and generates a predicted signal of each prediction block in the target encoded block. The predicted signal thus generated is output via line L103 to the adder 205.

[0090] The quantized transform coefficient decoder 207 entropy-decodes the encoded data of quantized transform coefficients of the residual signal in the target encoded block and outputs the result via line L207 to the inverse quantizer 203.

[0091] The inverse quantizer 203 performs inverse quantization of the residual signal information of the target encoded block fed via line L207.

The inverse transformer 204 performs an inverse discrete cosine transform of the inversely quantized data.

[0092] The adder 205 adds the predicted signal generated by the predicted signal generator 103, to the residual signal restored by the inverse quantizer 203 and the inverse transformer 204, and outputs a reconstructed pixel signal of the target encoded block via line L205 to the output terminal 206 and the frame memory 104. The output terminal 206 outputs the signal to the outside of the decoding device 200 (e.g., to a display).

[0093] The frame memory 104 stores the reconstructed image output from the adder 205, as a reference picture which is stored as a reconstructed image for reference for the next decoding processing.

[0094] Fig. 7 is a flowchart of the prediction information decoder 208 to implement the processing of Fig. 3.

[0095] First, the prediction information decoder 208 decodes the prediction block partitioning type of the target encoded block and stores it in the prediction information memory 115. At the same time, the prediction information decoder 208 sets the number N of prediction blocks in the target encoded block, based on the decoded prediction block partitioning type, and resets the target prediction block number i to 0 (step S251). Next, the prediction information decoder 208 determines whether a target prediction block is a prediction block to be decoded last in the target encoded block and whether the number of prediction blocks in the target encoded block is not less than 2 (step S252). For example, in the case of N=2, the determination is yes with i=1 and the processing proceeds to step S258. In the case of N=4 ((D)

of Fig. 20), the determination is yes with i=3. When the determination is no, the processing proceeds to step S253. In Fig. 3, the processing proceeds to step S253 when the target prediction block is the prediction block T1; the processing proceeds to step S258 when the target prediction block is the prediction block T2.

[0096] In step S253, the merging identification information is decoded. When the merging identification information is yes (merge_flag=1) herein, the merging identification information indicates that the predicted signal of the target prediction block is to be generated using a candidate for motion information. On the other hand, when the merging identification information is no (merge_flag=0), the predicted signal of the target prediction block is generated using the decoded motion information. In next step S254, the prediction information decoder 208 determines whether the merging identification information indicates decoding of motion information, i.e., whether the value of merge_flag is 0. When the decoded value of merge_flag is 0, the prediction information decoder 208 decodes the motion information for generation of the predicted signal of the target prediction block (step S257) and then the processing proceeds to step S267. When the value of merge_flag is 1, the prediction information decoder 208 determines in step S266 whether the number of candidates for motion information to be used in block merging is 2, and when the number of candidates is 2, the merging block selection information is decoded and the processing proceeds to step S256 (step S255). When the number of candidates for motion information to be used in block merging of the target prediction block is 1, the processing proceeds to step S256. In

step S256, when the number of candidates for motion information is 1, the prediction information decoder 208 determines the motion information thereof as the motion information of the target prediction block. When the number of candidates for motion information is 2, the prediction information decoder 208 determines the motion information of the neighboring block indicated by the merging block selection information, as the motion information of the target prediction block.

[0097] In step S258, the prediction information decoder 208 determines whether every piece of already-decoded motion information of the target encoded block coincides with the motion information of a neighboring block that does not belong to the target encoded block. The description of this step S258 means that, in the case of N=2, the motion information of the prediction block T1 shown in Fig. 3 is compared with the motion information of the neighboring block D. Furthermore, the description of this step S258 means that, in the case of N=4 ((D) of Fig. 20), the target prediction block is the lower right partitioned block and the pieces of motion information of the three other prediction blocks (upper left, upper right, and lower left) are compared with each other. When the determination is yes (or when the pieces of motion information in comparison are coincident), the number of candidates for motion information to be used in block merging of the target prediction block is 0, as shown in the example of (B) in Fig. 3, the prediction information decoder 208 decodes the motion information to be used for generation of the predicted signal of the target prediction block, without decoding the block merging information, and the processing proceeds to step S267 (step S262). On the other hand,

when the determination is no (or when the pieces of motion information in comparison are not coincident), the processing proceeds to step S265. In the case of N=4, the pieces of motion information of the upper right and the lower left blocks in the target encoded block are those of neighboring blocks to the target prediction block. For this reason, application of block merging to the target prediction block (lower right) when coincidence of the pieces of motion information of the three prediction blocks (upper left, upper right, and lower left) means that the predicted signals of the four prediction blocks in the target encoded block are generated all from the same motion information. For this reason, in the case where N=4 and where the pieces of motion information of the three prediction blocks (upper left, upper right, and lower left) are identical with each other, the number of candidates for motion information of the target prediction block (lower right) is set to 0.

[0098] In step S265, the prediction information decoder 208 determines whether the prediction block partitioning type of the target encoded block is a bisectional type, and if the determination is no, the processing proceeds to step S253 (the description hereinafter is omitted). When the determination in step S265 is yes, the processing proceeds to step S259, in which the prediction information decoder 208 decodes the merging identification information. In this case, as in the example of (A) of Fig. 3, the number of candidates for motion information to be used in block merging of the target prediction block is 1, and therefore there is no need for decoding of the merging block selection information.

[0099] In next step S260, the prediction information decoder 208 determines whether the merging identification information indicates decoding of motion information, i.e., whether the value of merge_flag is 0. When the decoded value of merge_flag is 0, the prediction information decoder 208 decodes the motion information for generation of the predicted signal of the target prediction block (step S262) and the processing proceeds to step S267. When the value of merge_flag is 1, the processing proceeds to step S261. In step S261, since the number of candidates for motion information is 1, as shown in (A) of Fig. 3, the prediction information decoder 208 determines the motion information of the neighboring block D as the motion information of the target prediction block and the processing proceeds to step S267.

[0100] In step S267, the motion information of the restored prediction block is stored into the prediction information memory 115. Subsequently, in step S263, the prediction information decoder 208 determines whether the decoding is complete for all the prediction blocks in the target encoded block (whether i=N−1); when i=N−1, this prediction information decoding processing of the target encoded block is terminated; when i < N−1, the number i is updated in step S264 (i=i+1) and the processing returns to step S252 to perform the motion information decoding processing of the next prediction block.

[0101] Next, an image predictive decoding method in the image predictive decoding device 200 shown in Fig. 6 will be described using Fig. 8. First, compressed data is input through the input terminal 201 (step S201). Then the data analyzer 202 performs data analysis of the compressed data to extract the encoded data of the prediction

information and the quantized transform coefficients about a target region of a decoding target. The prediction information is decoded by the prediction information decoder 208 (S203).

[0102] Thereafter, based on the restored prediction information, the predicted signal generator 103 generates the predicted signal of the target encoded block (S204).

[0103] The quantized transform coefficients decoded by the quantized transform coefficient decoder 207 are subjected to the inverse quantization in the inverse quantizer 203 and to the inverse transformation in the inverse transformer 204, to generate a reconstructed residual signal (S205). Then the generated predicted signal is added to the reconstructed residual signal to generate a reconstructed signal, and this reconstructed signal is stored into the frame memory 104, for reconstruction of the next target encoded block (step S206). If there is next compressed data, the processes of S204 to S206 are repeatedly carried out (S207) to process the entire data to the last.

[0104] The above described the examples in which the number of neighboring blocks to the prediction block is not more than 2, and the below will focus attention on situations in which the number of neighboring blocks in contact with upper and left block boundaries to a prediction block is not less than 3.

[0105] The example of Fig. 3 concerned the case where there were two neighboring blocks in contact with a prediction block, but there are situations in which a prediction block is in contact with two or more neighboring blocks, depending upon combinations of the prediction

block partitioning types of an encoding block and neighboring blocks thereto. Fig. 9 shows an example where three neighboring blocks are in contact with a prediction block. The block 301 in Fig. 20 will be described as an example herein, but the same description also applies to the blocks 302, 304, 305, 306, and 307.

[0106] In (A) and (B) of Fig. 9, a target encoded block 400 has two prediction blocks resulting from vertical bisection of the block 400, whereas a block 401 in contact with the left side of the prediction block T1 is horizontally bisected (or partitioned into two upper and lower blocks). For this reason, the prediction block T1 is in contact with three neighboring blocks A, B, and C. In this case, when it is preliminarily determined on the encoding side and on the decoding side that neighboring blocks are represented by two neighboring blocks A and B in contact with the upper left corner of the target prediction block, the number of neighboring blocks is always limited to 2 and therefore the technique described above is applicable.

[0107] On the other hand, it is also possible to employ a technique of virtually horizontally bisecting the prediction block T1 according to the prediction block partitioning type of the neighboring block 401, as shown in (B) of Fig. 9. In this case, the target prediction block T1 is partitioned into blocks T1a and T1b and the predicted signal of block T1a and the predicted signal of T1b are generated using two pieces of motion information belonging to the neighboring blocks A and C, respectively.

[0108] On this occasion, the merging block selection information can be efficiently encoded without change in the configuration of the block

merging information, by such a rule that selected candidates for merging block selection information are two pieces of the motion information of the neighboring block B in (A) of Fig. 9 and the combination of pieces of motion information of the neighboring blocks A and C in (B) of Fig. 9.

[0109] On the other hand, in the case where either of (A) of Fig. 9 and (B) of Fig. 9 is identified by the merging block selection information and where (B) of Fig. 9 is selected, it is also possible to adopt a method of further transmitting second merging identification information for each virtual block and identifying generation of the predicted signal of the virtual block based on the motion information of the neighboring block, or encoding/decoding of the motion information.

[0110] It is also possible to adopt a method without partition of the prediction block T1 in which selected candidates for the merging block selection information in the prediction block T1 are three pieces of motion information of neighboring blocks A, B, and C and in which the motion information to be used in generation of the predicted signal of T1 is selected from the three pieces of information, but changes below are needed in this case.

1. A flow of "acquiring the prediction block partitioning type of the neighboring block and deriving the number of blocks neighboring the prediction block" is added prior to step S164 in Fig. 4 and step S266 in Fig. 7.

2. Step S164 in Fig. 4 and step S266 in Fig. 7 are changed to "Are there two or more pieces of motion information of selected candidates?"

3. The merging block selection information is extended to information

to select one out of three or more candidates.

[0111] This processing of block merging shown in (A) and (B) of Fig. 9 can be implemented by extending step S256 in Fig. 7 to the processing shown in Fig. 11. First in step S256a, the prediction block partitioning type of an encoded block in contact with the target prediction block is acquired. In next step S256b, the number M of prediction blocks neighboring block boundaries indicated by the decoded merging block selection information is derived from the acquired prediction block partitioning type. For example, in the case of (B) of Fig. 9, M=2. Furthermore, in step S256c, it is determined whether the value of M is larger than 1 (M > 1). In the case of M > 1, the target prediction block is partitioned into M virtual blocks, and pieces of motion information of M neighboring blocks are set to the M partitioned virtual blocks (it can also be contemplated that merging identification information is additionally sent for each virtual block and it is determined whether the motion information is to be decoded). In the case of M=1, the motion information of a neighboring block serving as a candidate for block merging is set to the motion information of the target prediction block.

[0112] According to Figs. 7 and 11 as described above, the selection of candidate for motion information in the example of Fig. 9 is carried out based on the pieces of information below.

1) The number of already-encoded/already-decoded prediction blocks in the target encoded block

2) The prediction block partitioning type of the target encoded block

3) The prediction block partitioning type of the block neighboring the target prediction block

In this manner, the information of 3), which is not used in the selection of candidate for motion information in the example of Fig. 3, is used in the cases where there are three or more candidates for motion information.

[0113] (C) of Fig. 9 shows an example in which the block neighboring the left side of the prediction block 400 is asymmetrically bisected. In this case, it is also possible to adopt the technique of virtually horizontally bisecting the prediction block T1 according to the prediction block partitioning type of the neighboring block 401 (into blocks T1a and T1b). Namely, the predicted signal of the target prediction block T1 can be generated using a combination of pieces of motion information of neighboring blocks A and C in (C) of Fig. 9 as candidates for motion information for block merging of the prediction block T1.

[0114] In cases where the prediction block partitioning type of the encoding block is a type in which the number of prediction block is 1 like the block 300 in Fig. 20, as shown in (D) to (F) of Fig. 9, it is also possible to apply the technique of virtually horizontally partitioning the prediction block T1 (block 400) according to the prediction block partitioning type of neighboring block 401 (partition into a plurality of blocks arranged in the vertical direction), and generating the predicted signal for each block. Furthermore, in cases where the neighboring block 402 is vertically partitioned (into a plurality of blocks arranged in the horizontal direction), which are not shown, it is possible to apply a technique of virtually vertically partitioning the prediction block T1 (block 400) according to the prediction block partitioning type of the

neighboring block 402 and generating the predicted signal for each block.

[0115] In cases where a block neighboring the prediction block includes an intra-picture predicted block (intra), it is also possible to apply the technique of virtually partitioning the prediction block and generating the predicted signal, by preliminarily determining rules. (A) to (F) of Fig. 10 show examples in which an intra-picture predicted block (intra) is included in a plurality of neighboring blocks A, C, E, and G in contact with the left side of the prediction block. Based on the prediction block partitioning type of the neighboring block and the prediction mode (inter-picture/intra-picture prediction) in the prediction information, the intra-picture predicted block in the neighboring block is virtually integrated with an inter-picture predicted block with motion information (thick lines in the drawing). In these examples, an intra-picture predicted block is virtually integrated with an inter-picture predicted block which is closer to the upper left corner of the neighboring block and which is closest to the intra-picture block. As a consequence, the prediction block T1 is virtually partitioned according to the number of inter-picture predicted blocks in the neighboring block, as shown in (A) to (F) of Fig. 10. In this manner, even in the cases where the neighboring block includes an intra-picture predicted block (intra), the generation of the predicted signal by block merging can be carried out using the motion information of the inter-picture predicted block in the neighboring block.

[0116] There are no restrictions on the rules of integrating the intra-picture predicted block with the inter-picture predicted block in the

neighboring block. It can be contemplated that a plurality of rules as described above are prepared and a rule is selected for each frame or for each slice to implement encoding.

In this case, the selection of a candidate for motion information is carried out based on pieces of information below.

1) The number of already-encoded/already-decoded prediction blocks in the target encoding/encoded block

2) The prediction block partitioning type of the target encoding/encoded block

3) The prediction block partitioning type of the block neighboring the target prediction block

5) The prediction mode (intra-picture prediction/inter-picture prediction) of the block neighboring the target prediction block

[0117] Fig. 12 shows examples in which the encoding block 400 and neighboring block 402 are similarly vertically bisected but their partition shapes are different. In these examples, the prediction block T1 (block including blocks T1a and T1b) in (A) of Fig. 12 and the prediction block T2 (block including blocks T2a and T2b) in (B) of Fig. 12 also have three neighboring blocks. For T1 in (A) of Fig. 12, the processing flow of Fig. 11 is applied to step S256 in Fig. 7, whereby it becomes feasible to execute the block merging by setting pieces of motion information of blocks Ba and Bb to respective blocks T1a and T1b resulting from virtual vertical bisection of the prediction block T1. For T2 in (B) of Fig. 12, a processing flow of Fig. 13 described below is applied to step S261 in Fig. 7, whereby it becomes feasible to execute the block merging by setting pieces of motion information of blocks Ba

and Bb to respective blocks T2a and T2b resulting from virtual vertical bisection of the prediction block T2. On this occasion, it is also possible to adopt the method of transmitting second merging identification information for each virtual block and identifying either generation of the predicted signal of the virtual block based on the motion information of the neighboring block or encoding/decoding of motion information.

[0118] It is also possible to adopt a method in which the prediction block T2 is not partitioned, two pieces of motion information of block Ba and block Bb are defined as candidates for motion information to be used in block merging of the prediction block T2, and one of the pieces of motion information of block Ba and block Bb is selected as motion information to be used in generation of the predicted signal of T2, but in that case, it is necessary to extend the flow of Fig. 7 as described below.

1. A flow of "acquiring the prediction block partitioning type of the neighboring block and deriving the number of blocks neighboring the prediction block" is added after step S158 in Fig. 4 and after step S259 in Fig. 7.

2. Step S159 in Fig. 4 and step S260 in Fig. 7 are changed to "Are there two or more pieces of motion information of selected candidates?"

3. A step of performing encoding/decoding of block selection information is added after step S159 in Fig. 4 and after step S260 in Fig. 7.

[0119] The flow of Fig. 13 will be described below. In Fig. 13, first in step S261a, the prediction block partitioning type of the encoded block in contact with the target prediction block is acquired. In next step

S261b, the number M of prediction blocks neighboring block boundaries in contact with the neighboring block not belonging to the target encoded block is derived from the acquired prediction block partitioning type. For example, in the case shown in (B) of Fig. 12, M=2. Furthermore, it is determined in step S261c whether the value of M is larger than 1 (M > 1). In the case of M > 1, the target prediction block is partitioned into M virtual blocks and pieces of motion information of M neighboring blocks are set to the M partitioned virtual blocks (it is also possible to additionally send the merging identification information for each virtual block and determine whether the motion information is to be decoded). In the case of M=1, the motion information of the neighboring block as a candidate for block merging is set as the motion information of the target prediction block.

[0120] According to Figs. 12 and 13 as described above, the selection of a candidate for motion information in the example of Fig. 11 is carried out based on the pieces of information below.

1) The number of already-encoded/already-decoded prediction blocks in the target encoding/encoded block

2) The prediction block partitioning type of the target encoding/encoded block

3) The prediction block partitioning type of the block neighboring the target prediction block

It should be noted that although Fig. 11 described the example of vertical partition, the same processing is also applicable to examples of horizontal partition (partition into a plurality of blocks arranged in the vertical direction) like blocks 306 and 307 in Fig. 20.

[0121] It is possible to adopt further modifications described below.

[0122] (Candidates for Motion Information)

[0123] In the above description the pieces of motion information of blocks in contact with the upper side and the left side of the prediction block were defined as candidates for block merging, but it is also possible to set a limitation based on the prediction block partitioning types of the target encoding/encoded block and neighboring blocks, as shown in (A) and (B) of Fig. 14 and (A) of Fig. 15. (A) and (B) of Fig. 14 show examples where there are two neighboring blocks and where the motion information of neighboring blocks on the side in contact with two or more neighboring blocks out of the upper side and the left side of the prediction block is excluded from candidates for block merging. In this case, there is no need for encoding of the merging block selection information, which can reduce additional information. The candidates for motion information to be used in block merging of the prediction block T1 in (A) of Fig. 14 and the prediction block T1 in (B) of Fig. 14 are determined to be pieces of motion information of block B and block A, respectively.

[0124] (A) of Fig. 15 shows a technique of automatically selecting candidates for motion information to be used in block merging of prediction blocks T1 and T2, based on the prediction block partitioning type of the target encoding/encoded block.

[0125] (B) of Fig. 15 shows an example in which the prediction block to which the block merging is applied is limited according to the prediction block partitioning type of the target encoding block and the number of already-encoded/already-decoded blocks in the target

encoding block. In the example shown in Fig. 3, when the motion information of block T1 is coincident with that of the block D, the motion information of block D is excluded from candidates for motion information to be used block merging of block T2; whereas, in the case shown in (A) of Fig. 15, without comparison between the motion information of block T1 and the motion information of block D, the block D is excluded from candidates for block merging, based on the number of already-encoded/decoded prediction blocks in the target encoding/encoded block. In this manner, the prediction block to which block merging is applied may be limited by the number of motion vectors to be encoded in the target encoding block.

[0126] Furthermore, it is also possible to place a limitation, according to block sizes of two neighboring blocks in contact with the upper left corner of the prediction block and the block size of the prediction block. For example, when the size of the right side of the neighboring block in contact with the left side of the target prediction block is smaller than a preset size (e.g., a half or a quarter of the length of the left side of the prediction block), the motion information of the neighboring block may be excluded from candidates for block merging of the target prediction block.

[0127] When the limitation is set on candidates for motion information as described above, the code amount of block merging information can be reduced.

[0128] (Selection of Candidate for Motion Information)

[0129] The selection of candidate for motion information is carried out based on the pieces of information below, but a method of using the

information is not limited to the methods described above. The means to select the candidates for motion information using these pieces of information can be implemented by the configurations of Fig. 1 and Fig. 6.

1) The number of already-encoded/already-decoded prediction blocks in the target encoding/encoded block

2) The prediction block partitioning type of the target encoding/encoded block

3) The prediction block partitioning type of the block neighboring the target prediction block

4) The motion information of already-encoded/already-decoded prediction blocks in the target encoding/encoded block

5) The motion information and prediction mode (intra-picture prediction/inter-picture prediction) of the block neighboring the target prediction block

[0130] (Encoding of Prediction Block)

[0131] In the above-described description the encoding/decoding of the prediction blocks in the encoding block is carried out in a raster scan order, but the above-described selection of candidate for motion information to be used in the block merging is also applicable to cases where the prediction blocks are encoded/decoded in any order. For example, in the example of Fig. 3, where the prediction block T2 of the target encoding/encoded block 400 is first encoded/decoded, the motion vector of the prediction block T2 is not included as a candidate for motion information to be used in block merging of the prediction block T1.

[0132] (Shape of Block)

[0133] In the above-described description the partial regions in the encoding block were always rectangular, but they may be any shape. In this case, shape information may be included in the prediction information of the encoding block.

[0134] (Transformer and Inverse Transformer)

[0135] The transformation process of the residual signal may be carried out in a fixed block size, or the transformation process may be carried out after a target region is sub-partitioned according to partial regions.

[0136] (Prediction Information)

[0137] In the above-described description the generation method of the predicted signal was described as inter-picture prediction (prediction using the motion vector and reference frame information), but the generation method of the predicted signal is not limited to this. The aforementioned predicted signal generation process is also applicable to the intra-picture prediction and the prediction method including luminance compensation or the like. In this case, the prediction information contains mode information, luminance compensation parameters, and so on.

[0138] In Fig. 10 the intra-picture predicted block in the neighboring block is virtually integrated with the inter-picture predicted block, but it is also possible to adopt a method in which the prediction block is virtually partitioned, regardless of the prediction mode of the neighboring block, and partial signals in the prediction block are predicted by intra-picture prediction.

[0139] (Color Signal)

[0140] The foregoing description contains no particular description of color format, but the predicted signal generation process may also be carried out for color signal or color-difference signal, independently of luminance signal. The predicted signal generation process may also be carried out in synchronism with the process of luminance signal.

[0141] (Block Noise Removal Process)

[0142] Although the foregoing description states nothing, the reconstructed picture may be subjected to a block noise removal process and in that case, it is preferable to perform the noise removal process for boundary portions of partial regions. In the cases where the prediction block is virtually partitioned in the examples shown in Figs. 9, 10, and 12, the block noise removal process may also be applied to a boundary between virtually partitioned blocks.

[0143] The image predictive encoding method and the image predictive decoding method according to the embodiments of the present invention can also be provided as stored in the form of programs in a recording medium. Examples of recording media include recording media such as floppy disks (registered trademark), CD-ROMs, DVDs, or ROMs, or semiconductor memories, or the like.

[0144] Fig. 16 is a block diagram showing modules of a program which can execute the image predictive encoding method. The image predictive encoding program P100 is provided with block partition module P101, motion information estimation module P102, 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 partitioning type selection module P112, prediction information storage module P113, and prediction information encoding module P114. The functions implemented upon execution of the respective modules by a computer are the same as the functions of the aforementioned image predictive encoding device 100. Namely, the block partition module P101, motion information estimation module P102, 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 partitioning type selection module P112, prediction information storage module P113, and prediction information encoding module P114 cause the computer to execute the same functions as the block partitioner 102, motion information estimator 114, predicted signal generator 103, frame memory 104, subtracter 105, transformer 106, quantizer 107, inverse quantizer 108, inverse transformer 109, adder 110, quantized transform coefficient encoder 111, prediction block partitioning type selector 113, prediction information memory 115, and prediction information encoder 116, respectively.

[0145] Fig. 17 is a block diagram showing modules of a program which can execute the image predictive decoding method. The image predictive decoding program P200 is provided with quantized transform coefficient decoding module P201, prediction information decoding module P202, prediction information storage module P113, inverse quantization module P206, inverse transform module P207, addition

module P208, predicted signal generation module P103, and storage module P104.

[0146] The functions implemented upon execution of the respective modules are the same as those of the respective components of the aforementioned image predictive decoding device 200. Namely, the quantized transform coefficient decoding module P201, prediction information decoding module P202, prediction information storage module P113, inverse quantization module P206, inverse transform module P207, addition module P208, predicted signal generation module P103, and storage module P104 cause the computer to execute the same functions as the quantized transform coefficient decoder 207, prediction information decoder 208, prediction information memory 115, inverse quantizer 203, inverse transformer 204, adder 205, predicted signal generator 103, and frame memory 104, respectively.

[0147] The image predictive encoding program P100 or the image predictive decoding program P200 configured as described above is stored in a recording medium SM and executed by the computer described below.

[0148] Fig. 18 is a drawing showing a hardware configuration of the computer for executing the programs recorded in the recording medium and Fig. 19 a perspective view of the computer for executing the programs stored in the recording medium. Equipment to execute the programs stored in the 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.

[0149] As shown in Fig. 19, the computer C10 is provided with a reading device C12 such as a floppy disk drive unit, a CD-ROM drive unit, or a DVD drive unit, a working memory (RAM) C14 in which an operating system is resident, a memory C16 to store programs stored in the recording medium SM, a monitor device C18 such as a display, a mouse C20 and a keyboard C22 as input devices, a communication device C24 for transmission and reception of data and others, and a CPU C26 to control execution of programs. When the recording medium SM is put into the reading device C12, the computer C10 becomes accessible to the image predictive encoding/decoding program stored in the recording medium SM, through the reading device C12, and becomes able to operate as the image encoding device or the image decoding device according to the embodiment of the present invention, based on the image encoding or decoding program.

[0150] As shown in Fig. 18, the image predictive encoding program and the image decoding program may be provided in the form of computer data signal CW superimposed on a carrier wave, through a network. In this case, the computer C10 becomes able to execute the image predictive encoding program or the image predictive decoding program after the image predictive encoding program or the image decoding program received by the communication device C24 is stored in the memory C16.

[0151] Still another embodiment will be described below. Fig. 24 is a drawing schematically showing a configuration of a video encoding device according to an embodiment. The video encoding device 10 shown in Fig. 24 is provided with block partitioner 501, sub-partition

generator 502, frame memory 503, motion detector 504, predicted signal generator 505, motion predictor 506, subtracter 507, residual signal generator 508, transformer 509, quantizer 510, inverse quantizer 511, inverse transformer 512, adder 513, and entropy encoder 514. An input image signal (video signal) fed into this video encoding device 10 is composed of a time sequence of image signals of frame units (hereinafter referred to as frame image signals).

[0152] The block partitioner 501 sequentially selects frame image signals, or input images serving as encoding targets from the input image signal fed via line L501. The block partitioner 501 partitions an input image into a plurality of partitions, or blocks. The block partitioner 501 sequentially selects the plurality of blocks as encoding target blocks and outputs a pixel signal of each of the target blocks (hereinafter referred to as a target block signal) via line L502.

[0153] In the video encoding device 10, encoding processing described below is carried out in block units. The block partitioner 501 may partition, for example, an input image into a plurality of blocks each having 8×8 pixels. However, the blocks may be any size and shape. The blocks may be, for example, blocks each having 32×16 pixels or blocks each consisting of 16×64 pixels.

[0154] The sub-partition generator 502 partitions a target block fed via line L502, into a plurality of sub-partitions. Fig. 25 is a drawing for explaining the generation of sub-partitions. As shown in Fig. 25, the sub-partition generator 502 partitions the target block P into two sub-partitions SP1 and SP2 by a straight line Ln expressed by the linear expression of formula (1).

$$y = mx + k \qquad (1)$$

[0155] For example, the sub-partition generator 502 can be configured as follows: with changes of the parameters m and k, it obtains a predicted signal of the sub-partition SP1 and a predicted signal of the sub-partition SP2, and determines m and k that minimize an error between the predicted signal of the sub-partition SP1 and an image signal of the sub-partition SP1 and an error between the predicted signal of the sub-partition SP2 and an image signal of the sub-partition SP2, as parameters of the straight line Ln.

[0156] The sub-partition generator 502 outputs the parameters m and k in formula (1) thus determined, as shape information to specify the shapes of the sub-partitions in the target block P, i.e., as shape information to specify the shapes of the first sub-partition SP1 and the second sub-partition SP2, via line L504.

[0157] The linear expression expressing the straight line Ln may be any one. For example, the straight line Ln may be one expressed by formula (2).

$$y = -x/\tan\theta + \rho/\sin\theta \qquad (2)$$

In this case, the shape information is θ and ρ.

[0158] The shape information may be information indicative of two arbitrary points that the straight line Ln passes, e.g., intersections between the straight line and boundaries of the block P. The block does not always have to be partitioned by a straight line, but the sub-partitions may be generated based on a pattern selected from a plurality of patterns prepared in advance. In this case, information such as an index to specify the selected pattern can be used as shape

information.

[0159] In the description below, coordinates are set with an origin at the uppermost and leftmost position of the target block, a sub-partition including the uppermost and leftmost pixel in the target block P is defined as a first sub-partition, and the other as a second sub-partition. It is, however, noted that any defining method is applicable herein: for example, a sub-partition that does not include the position of the center in the target block can be defined as a first sub-partition, and the other as a second sub-partition. In this case, the shape information may be intersection information of block boundaries or pattern identification information.

[0160] The frame memory 503 stores signals of previously-reconstructed images fed via line L505, i.e., frame image signals having been encoded in the past (which will be referred to hereinafter as reference frame image signals). The frame memory 503 outputs the reference frame images signal via line L506.

[0161] The motion detector 504 receives the target block signal fed through line L502, the shape information of the block fed through line L504, and the reference frame image signals fed through line L506. The motion detector 504 searches image signals in a predetermined range of the reference frame image signals, for a signal similar to an image signal of a sub-partition serving as a processing target, and calculates a motion vector. This motion vector is a spatial displacement amount between a region in a reference frame image signal having a pixel signal similar to the image signal of the sub-partition serving as a processing target, and the target block. The

motion detector 504 outputs the motion vector thus calculated, via line L507.

[0162] The motion detector 504 may be configured, at the same time, to also detect a motion vector for a target block and to determine whether a predicted signal is to be generated for each of two sub-partitions resulting from partition of the target block. This determination may be such determination that if an error between the predicted signal of the target block and the image signal of the target block is smaller than errors between the predicted signals of two sub-partitions generated by partition of the target block, and the image signals of the two sub-partitions, the target block is not partitioned into the sub-partitions. When this determination is made, information indicative of the result of the determination is encoded as partition applicability information and the shape information may be encoded only if the partition applicability information indicates that a target block is to be partitioned into sub-partitions.

[0163] The predicted signal generator 505 generates the predicted signal of the image signal of the sub-partition serving as a processing target, based on the motion vector fed via line L507 and the block shape information fed via line L504, from the image signal in the predetermined range of the reference frame image signal fed via line L506.

[0164] The predicted signal generator 505 combines the predicted signals of the respective sub-partitions in the target block to generate the predicted signal of the target block. The predicted signal generator 505 outputs the predicted signal thus generated, via line L508. The

predicted signal may be generated by intra-picture prediction, instead of the inter-picture prediction.

[0165] The motion predictor 506 generates a motion vector predictor of a processing target sub-partition in a target block, based on the block shape information fed via line L504, the motion vector fed via line L507, and a motion vector of a block prior in order to the processing target sub-partition or a motion vector of an already-processed partial region that is a sub-partition. The motion predictor 506 outputs the motion vector predictor thus generated, via line L509.

[0166] The motion predictor 506 may select one motion vector predictor out of a plurality of candidates for motion vector predictor. In this case, the motion predictor 506 also outputs indication information to specify the selected motion vector predictor, via line L510. If the candidates for motion vector predictor of the processing target sub-partition are narrowed down to one according to a predetermined rule shared with the decoder side, the output of the indication information may be omitted.

[0167] The subtracter 507 subtracts the motion vector predictor fed via line L509, from the motion vector of the processing target sub-partition fed via line L507, to generate a differential motion vector. The subtracter 507 outputs the differential motion vector thus generated, via line L511.

[0168] The residual signal generator 508 subtracts the predicted signal of the target block fed via line L508, from the target block signal fed via line L502, to generate a residual signal. The residual signal generator 508 outputs the residual signal thus generated, via line L512.

[0169] The transformer 509 performs orthogonal transformation of the residual signal fed via line L512, to generate transform coefficients. The transformer 509 outputs the transform coefficients thus generated, via line L513. This orthogonal transformation may be performed, for example, by DCT. However, the transformation used by the transformer 509 may be any transformation.

[0170] The quantizer 510 quantizes the transform coefficients fed via line L513, to generate quantized transform coefficients. The quantizer 510 outputs the quantized transform coefficients thus generated, via line L514.

[0171] The inverse quantizer 511 performs inverse quantization of the quantized transform coefficients fed via line L514, to generate inversely-quantized transform coefficients. The inverse quantizer 511 outputs the inversely-quantized transform coefficients thus generated, via line L515.

[0172] The inverse transformer 512 performs inverse orthogonal transformation of the inversely-quantized transform coefficients fed via line L515, to generate a reconstructed residual signal. The inverse transformer 512 outputs the reconstructed residual signal thus generated, via line L516. The inverse transformation used by the inverse transformer 512 is a process symmetric with the transformation of the transformer 509.

[0173] The transformation is not always essential, and the video encoding device does not always have to be provided with the transformer 509 and the inverse transformer 512. Likewise, the quantization is not always essential, and the video encoding device does

not always have to be provided with the quantizer 510 and the inverse quantizer 511.

[0174] The adder 513 adds the reconstructed residual signal input via line L516, to the predicted signal of the target block fed via line L508, to generate a reconstructed image signal. The adder 513 outputs the reconstructed image signal as a previously-reconstructed image signal via line L505.

[0175] The entropy encoder 514 encodes the quantized transform coefficients fed via line L514, the shape information of the target block fed via line L504, the indication information of the motion vector predictor fed via line L510, and the differential motion vector fed via line L511. The entropy encoder 514 multiplexes codes generated by encoding, to generate a compressed stream, and then outputs the compressed stream through line L517.

[0176] The entropy encoder 514 may use any encoding method such as the arithmetic coding or the run-length coding. The entropy encoder 514 may adaptively determine an occurrence probability in arithmetic coding of the indication information of the motion vector predictor fed via line L510, based on the shape information of the target block fed via line L504. For example, the entropy encoder 514 may set a high value as an occurrence probability of the indication information to indicate a motion vector of a partial region in contact with a processing target sub-partition.

[0177] Fig. 26 is a drawing showing a configuration of the motion predictor according to an embodiment. As shown in Fig. 26, the motion predictor 506 has a motion vector memory 5061, a motion

reference candidate generator 5062, and a motion vector predictor generator 5063.

[0178] The motion vector memory 5061 stores motion vectors of previously-processed partial regions and outputs the previously-encoded motion vectors via line L5061, for derivation of a motion vector predictor of a processing target sub-partition.

[0179] The motion reference candidate generator 5062 generates candidates for motion vector predictor from the motion vectors of the partial regions fed via line L5061, by a below-described method, based on the shape information fed via line L504. The motion reference candidate generator 5062 outputs the candidates for motion vector predictor thus generated, via line L5062.

[0180] The motion vector predictor generator 5063 selects a candidate that minimizes the difference from the motion vector of the processing target sub-partition, from the candidates for motion vector predictor fed via line L5062. The motion vector predictor generator 5063 outputs the selected candidate as a motion vector predictor via line L509. It also outputs the indication information to specify the selected candidate, via line L510.

[0181] If the number of candidates generated in the motion reference candidate generator is limited to one, the output of the indication information can be omitted. There are no restrictions on a method to limit the number of candidates to one, but any method can be applied, for example, such as a method of using an intermediate value of three candidates, a method of using an average of two candidates, and a method of defining a priority order for selection of one out of a plurality

of candidates.

[0182] The below will describe the operation of the video encoding device 10 and also describe a video encoding method according to an embodiment. Fig. 27 is a flowchart of the video encoding method according to an embodiment.

[0183] In an embodiment, as shown in Fig. 27, the block partitioner 501 first partitions an input image into a plurality of blocks, in step S501. In next step S502, the sub-partition generator 502 partitions a target block into a plurality of sub-partitions, as described above. The sub-partition generator 502 also generates the shape information as described above.

[0184] In step S503, the motion detector 504 then obtains a motion vector of a processing target sub-partition, as described above. In subsequent step S504, the predicted signal generator 505 generates a predicted signal of the target block, using the motion vectors of the respective sub-partitions in the target block and the reference frame image signals, as described above.

[0185] In step S505, the motion predictor 506 then obtains a motion vector predictor. Furthermore, the motion predictor 506 generates the indication information to specify a candidate selected from a plurality of candidates for motion vector predictor. The details of the process of this step S505 will be described later. In subsequent step S506, the subtracter 507 calculates the difference between the motion vector of each subblock and the motion vector predictor to generate a differential motion vector as described above.

[0186] In step S507, the residual signal generator 508 then obtains the

difference between the image signal of the target block and the predicted signal to generate a residual signal. In subsequent step S508, the transformer 509 performs orthogonal transformation of the residual signal to generate transform coefficients. In subsequent step S509, the quantizer 510 quantizes the transform coefficients to generate quantized transform coefficients. In subsequent step S510, the inverse quantizer 511 performs inverse quantization of the quantized transform coefficients to generate inversely-quantized transform coefficients. In subsequent step S511, the inverse transformer 512 performs inverse transformation of the inversely-quantized transform coefficients to generate a reconstructed residual signal.

[0187] In step S512, the adder 513 then adds the predicted signal of the target block to the reconstructed residual signal to generate a reconstructed image signal. In subsequent step S513, the reconstructed image signal is stored as a previously-reconstructed image signal in the frame memory 503.

[0188] In step S514, the entropy encoder 514 then encodes the quantized transform coefficients, the shape information of the target block, the indication information of the motion vector predictor, and the differential motion vector.

[0189] In next step S515, it is determined whether all blocks have been processed. If the processing is not completed for all the blocks, the processing from step S502 is continued for an unprocessed block as a target. On the other hand, if the processing is completed for all the blocks, the processing is terminated.

[0190] The operation of the motion predictor 506 will be described

below in more detail. Fig. 28 is a flowchart showing the process of the motion predictor according to an embodiment. The motion predictor 506 outputs the motion vector predictor (hereinafter PMV) and the indication information to specify PMV, according to the flowchart shown in Fig. 28.

[0191] In the process of the motion predictor 506, as shown in Fig. 28, the value of counter i is first set to 0, in step S505-1. It is assumed below that the process for the first sub-partition is carried out with i=0 and the process for the second sub-partition is carried out with i=1.

[0192] Next, step S505-2 is to generate candidates for PMV of a processing target sub-partition from motion vectors of previously-processed partial regions, according to a below-described method. The number of candidates for PMV is two in the present example. Namely, candidates for PMV may be set as follows: a motion vector of a previously-processed partial region located to the left of the processing target sub-partition and a motion vector of a previously-processed partial region located above the processing target sub-partition are set as candidates for the motion vector predictor of the processing target sub-partition. In step S505-2, the number of generated candidates is set in Ncand.

[0193] Next, in step S505-3, it is determined whether NCand is "0." When NCand is "0" (Yes), the processing proceeds to step S505-4. When NCand is not "0" (No), the processing proceeds to step S505-5.

[0194] In step S505-4, PMV is set to a zero vector and the processing proceeds to step S505-10. On this occasion, PMV may be set to a motion vector of a predetermined block, a motion vector of a partial

region processed immediately before the processing target sub-partition, or the like, instead of the zero vector.

[0195] In step S505-5, it is determined whether NCand is "1." When NCand is "1" (Yes), the processing proceeds to step S505-10. When NCand is not "1" (No), the processing proceeds to step S505-6.

[0196] In step S505-6, a PMV is selected from the candidates for PMV generated in step S505-2. The PMV to be selected can be a candidate that minimizes the difference from the motion vector of the processing target sub-partition.

[0197] Next, S505-7 is to determine whether the PMV selected in step S505-6 is a left candidate, i.e., a motion vector of the left partial region. When the PMV selected in step S505-6 is the left candidate (Yes), the processing proceeds to step S505-8. When the PMV selected in step S505-6 is not the left candidate (No), the processing proceeds to step S505-9.

[0198] In step S505-8, the indication information pmv_left_flag=1 to indicate that the PMV is the motion vector of the partial region located to the left of the processing target sub-partition is output. On the other hand, in step S505-9 the indication information pmv_left_flag=0 to indicate that the PMV is the motion vector of the partial region located above the processing target sub-partition is output.

[0199] Next, in step S505-10, the PMV remaining as a candidate is output. In subsequent step S505-11, "1" is added to the value of counter i.

[0200] Next, in step S505-12, it is determined whether the value of counter i is smaller than "2." When the value of counter i is smaller

than "2" (Yes), the processing proceeds to step S505-2. When the value of counter i is not smaller than "2" (No), the processing is terminated.

[0201] If step S505-2 is configured to limit the number of generated candidates to one, the steps S505-5, S505-6, S505-7, S505-8, and S505-9 can be omitted. There are no restrictions on this limiting method, but it is possible to adopt, for example, such a method as a method of using an intermediate value of three candidates, a method of using an average of two candidates, or a method of determining a priority order for selection of one out of a plurality of candidates, as described above in the description of the motion vector predictor generator 5063. In the configuration wherein the number of candidates generated in step S505-2 is limited to one, when NCand is not "0" in step S505-3 (No), the processing proceeds to step S505-10.

[0202] The method of generating candidates for the motion vector predictor of the processing target sub-partition in step S505-2 will be described below in more detail. Fig. 29 is a drawing showing an example of sub-partitions of a target block and surrounding partial regions.

[0203] The motion reference candidate generator 5062, as shown in Fig. 29, refers to partial region U1 and partial region L1 for the first sub-partition SP1 and, when each of the partial regions is one having been processed by inter-picture prediction, the motion reference candidate generator 5062 employs the motion vector of the partial region as a candidate for the motion vector predictor of the first sub-partition SP1. Similarly, the motion reference candidate generator

5062 refers to partial region U2 or partial region L2 for the second sub-partition to generate candidates for the motion vector predictor of the second sub-partition. The partial regions U1, L1, U2, and L2 herein are blocks or sub-partitions around the target block P and regions serving as units of generation of the predicted signal. The partial regions may be blocks prepared for generation of candidates for the motion vector predictor (e.g., blocks generated by partition in a single shape), irrespective of the units of generation of the predicted signal.

[0204] The partial region U1 is a partial region including a pixel Pi1(0,−1) neighboring above to the uppermost and leftmost pixel F(0,0) of the first sub-partition SP1, which is a previously-processed partial region in contact with the sub-partition SP1. The partial region L1 is a partial region including a pixel Pi2(−1,0) neighboring left to the uppermost and leftmost pixel F(0,0) of the first sub-partition SP1, which is a partial region in contact with the first sub-partition SP1. The partial region U2 is a partial region neighboring right to a partial region including a pixel Pi3(x1,−1), which is a partial region in contact with the x-axis. The partial region L2 is a partial region neighboring below to a partial region including a pixel Pi4(−1,y1), which is a partial region in contact with the y-axis.

[0205] The x-coordinate x1 of the pixel Pi3 and the y-coordinate y1 of the pixel Pi4 may be calculated by formula (3) and formula (4).

$$x1 = \mathrm{ceil}(-k/m) \qquad (3)$$

$$y1 = \mathrm{ceil}(k) \qquad (4)$$

Formulae (3) and (4) are formulae obtained by applying the function ceil(z) to values resulting from substitution of y=0 and x=0, respectively,

into the linear expression (1) to express an extension line Ln of a boundary as a partition between the first sub-partition SP1 and the second sub-partition SP2. The function ceil(z) is called a ceiling function, which is a function for deriving a minimum integer of not less than z, for real number z.

[0206] A floor function may be employed instead of the ceil function. The function floor(z) is called a floor function, which is a function for deriving a maximum integer of not more than z, for real number z.

[0207] Furthermore, x1 and y1 may be calculated by formulae (5) and (6).

$$x1 = \mathrm{ceil}((-1-k)/m) \qquad (5)$$

$$y1 = \mathrm{ceil}(-m+k) \qquad (6)$$

Formulae (5) and (6) are formulae obtained by applying the function ceil(z) to values resulting from substitution of y=−1 and x=−1, respectively, into formula (1).

[0208] Whether the partial regions U2 and L2 are existent is determined as described below. Conditions for existence of the partial region U2 are that it is in a picture and that formula (7) is satisfied. Conditions for existence of the partial region L2 are that it is in a picture and that formula (8) is satisfied.

$$0 < x1 \qquad (7)$$

$$0 < y1 \qquad (8)$$

[0209] When the condition of formula (7) is not satisfied, the partial region L2 exists between the second sub-partition SP2 and the partial region U2. In that case, the partial region U2 more distant from the second sub-partition SP2 than the partial region L2 closer to the second

sub-partition SP2 is less likely to have a motion vector close to that of the second sub-partition SP2. In this case, the motion vector of the partial region U2 can be excluded from candidates for the motion vector predictor by the condition of formula (7).

[0210] Likewise, when the condition of formula (8) is not satisfied, the partial region U2 exists between the second sub-partition SP2 and the partial region L2. In that case, the partial region L2 more distant from the second sub-partition SP2 than the partial region U2 closer to the second sub-partition SP2 is less likely to have a motion vector close to that of the second sub-partition SP2. In that case, the motion vector of the partial region U2 can be excluded from candidates for the motion vector predictor by the condition of formula (8).

[0211] In an example, the conditions defined by formulae (9) and (10) below may be used instead of the conditions of formulae (7) and (8).

$$0 < x1 < \text{blocksizeX} \qquad (9)$$

$$0 < y1 < \text{blocksizeY} \qquad (10)$$

Here, blocksizeX and blocksizeY are the number of horizontal pixels and the number of vertical pixels in the target block P. For example, when the target block P is a block of 8×8 pixels, blocksizeX=8 and blocksizeY=8.

[0212] By using the condition of formula (9) or formula (10), it is possible to exclude from candidates for the motion vector predictor, a motion vector of a partial region having no contact with the second sub-partition SP2, out of the partial region U2 and the partial region L2. This allows only candidates for the motion vector predictor with conceivably high prediction accuracy to be left.

[0213] When the partial regions U1, L1, U2, and L2 are set as described above, candidates for the motion vector predictor of each sub-partition are generated from motion vectors of previously-processed partial regions located on the same side with respect to the extension line of the boundary between sub-partitions.

[0214] As long as the candidates for the motion vector predictor of the sub-partition SP2 are generated from motion vectors of partial regions in the same domain as the sub-partition SP2 with respect to the extension line Ln of the boundary between the sub-partition SP2 and the other sub-partitions of the target block including the sub-partition SP2, the generation method of the motion vector predictor is not limited to that in the above-described embodiment. For example, the partial region U2 may be a partial region including the pixel Pi3 and the partial region L2 may be a partial region including the pixel Pi4.

[0215] A condition that the entire partial region is present in the same domain as the sub-partition SP2 with respect to the line Ln may be added as a condition for the motion vector of the partial region to be added to candidates for the motion vector predictor of the sub-partition SP2. In this case, it is possible to employ, for example, a method of inspecting positions of all the corners of the partial region.

[0216] Even if a partial region is not completely included in the same domain as a sub-partition with respect to the extension line, the motion vector of the partial region may be employed as a candidate for the motion vector predictor of the sub-partition. Fig. 30 is a drawing showing another example of sub-partitions of a target block and surrounding partial regions. As shown as an example in Fig. 30,

motion vectors of partial regions R_A, R_B, R_G, and R_E may be used as candidates for the motion vector predictor of the first sub-partition SP1. A motion vector predictor of partial region R_E may be added to candidates for the motion vector predictor of the second sub-partition SP2.

[0217] In the description about Figs. 28 and 29, the number of motion vectors serving as candidates for the motion vector predictor was at most two, but it is also possible to select two from the motion vectors obtained by any of the above-described conditions. For example, the motion vector of the partial region U2 shown in Fig. 29, and a motion vector of a partial region neighboring the partial region U2 may be selected as candidates for the motion vector predictor. Likewise, the motion vector of the partial region L2 and a motion vector of a partial region neighboring the partial region U2 may be selected as candidates for the motion vector predictor. Furthermore, three or more motion vectors may be selected as candidates for the motion vector predictor from the motion vectors specified by any of the above-described conditions. Moreover, an average or a median of a plurality of candidates for the motion vector predictor may be added to candidates for the motion vector predictor.

[0218] The block shape information can be used as a method for limiting the number of candidates for the motion vector predictor generated in step S505-2 in Fig. 28 to at most one. For example, out of previously-encoded partial regions in contact with a processing target sub-partition, a motion vector of a partial region with a maximum length of a portion in contact with the sub-partition may be added as a

candidate for the motion vector predictor. It is also possible to employ a motion vector of a previously-encoded partial region with a minimum shortest distance from a processing target sub-partition, as a candidate for the motion vector predictor of the sub-partition.

[0219] The above-described generation methods of candidates for the motion vector predictor can be applied to sub-partitions of any shape. Fig. 31 is a drawing showing further examples of sub-partitions of a target block and surrounding partial regions. (A) of Fig. 31 shows sub-partitions defined by a line Ln with a coordinate of y-intercept and a slope different from those of the line Ln shown in Fig. 29. (B) of Fig. 31 shows sub-partitions defined by a line Ln with a slope approximately symmetric with that of the line Ln with respect to the y-axis and with a coordinate of y-intercept different from that of the line Ln shown in Fig. 29. (C) of Fig. 31 shows sub-partitions defined by two lines Ln1 and Ln2. (D) of Fig. 31 shows sub-partitions defined by two lines Ln1 and Ln2 intersecting with each other. When the extension line of the boundary as shown in (A) to (D) of Fig. 31 is used as reference, the partial regions L2 and U2 with motion vectors that can be candidates for the motion vector predictor of the sub-partition SP2 can be specified by the aforementioned generation methods of candidates for the motion vector predictor.

[0220] It is noted that the sub-partitions are not limited only to those partitioned by a straight line. For example, in the case where the shapes of sub-partitions are selected out of predetermined patterns, a motion vector of a previously-encoded partial region belonging to the same domain as a processing target sub-partition with respect to an

extension line of a boundary between sub-partitions can be used as a candidate for the motion vector predictor. If patterns of sub-partition shapes are preliminarily defined, it is also possible to preliminarily determine a partial region with a motion vector to be adopted as a candidate for the motion vector predictor, for each shape pattern. The patterns may include the patterns to partition a target block into rectangular sub-partitions.

[0221] The aforementioned selection method of the motion vector predictor can also be applied as a method of selecting a motion vector in generation of a predicted signal of a processing target sub-partition using motion vectors of previously-encoded partial regions. Namely, the predicted signal of the processing target sub-partition may be generated using the motion vector predictor selected in step S505-2 in Fig. 28. In this case, there is no need for encoding of the differential motion vector and therefore the motion vector predictor output from the motion predictor 506 is not output to the subtracter 507 but to the predicted signal generator 505.

[0222] Furthermore, the video encoding device 10 may be configured to determine whether the differential motion vector is to be encoded, and to encode application information to specify the result of the determination. In this modification, the motion predictor 506 may include a function to switch output of the motion vector predictor either to the subtracter 507 or to the predicted signal generator 505, based on the application information.

[0223] In this modification, it is unfavorable that motion vectors of all the sub-partitions in a target block become identical with each other,

because the partition of the target block becomes meaningless. Namely, on the occasion of generating the candidates for motion vector of a processing target sub-partition in step S505-2 in Fig. 28, a motion vector of a previously-encoded sub-partition in the target block may be excluded from the candidates. For example, in the case where the target block is partitioned into two sub-partitions and where the motion vector of the first sub-partitions is first encoded, the motion vector of the first sub-partitions is excluded from the candidates for the motion vector predictor of the second sub-partitions. If the motion vector of the first sub-partitions is the same as that of the partial region U2, the motion vector of the partial region U2 does not have to be used in generation of the motion vector predictor of the second sub-partition.

[0224] If whether the differential motion vector is to be encoded is indicated, the occurrence probability in arithmetic coding of the aforementioned application information may be adaptively determined according to the sub-partition shape information. For example, the occurrence probability for the application information to indicate that the differential motion vector of the first sub-partition is not encoded can be set higher than that for the application information to indicate that the differential motion vector of the second sub-partition is not encoded. The reason for it is as follows: the second sub-partition may have no contact with any previously-encoded partial region, while the first sub-partition always has a contact with a previously-encoded partial region; therefore, the setting of the occurrence probabilities as described above may reduce a code amount of the application information.

[0225] The effect of one embodiment will be described with reference to Fig. 32 showing an example of partition of a target block into rectangular sub-partitions, for simplicity. In this example, the target block P is partitioned into a left sub-partition SP1 and a right sub-partition SP2 by a straight line Ln. In this example, a motion vector of the first sub-partition SP1 and a motion vector of a partial region R_B are candidates for the motion vector predictor of the second sub-partition SP2.

[0226] In the example shown in Fig. 32, if the predicted signal of the second sub-partition SP2 is generated using the motion vector of the first sub-partition SP1, the predicted signal of the first sub-partition SP1 and the predicted signal of the second sub-partition SP2 will be generated using the same motion vector, which makes the partition of the target block into two sub-partitions meaningless. For this reason, the predicted signal of the second sub-partition SP2 may be generated using the motion vector of the partial region R_B above the sub-partition SP2. In the example shown in Fig. 32, therefore, it is preliminarily determined between the encoding device and the decoding device that the predicted signal of the second sub-partition SP2 is to be generated using the motion vector of the partial region R_B, which reduces candidates for the motion vector predictor and which eliminates a need for transmitting indication information to indicate one motion vector predictor out of a plurality of candidates for the motion vector predictor.

[0227] Furthermore, a method for the video encoding device 10 to determine whether the differential motion vector needs to be encoded (in which the motion predictor 506 switches the output of the motion

vector predictor either to the subtracter 507 or to the predicted signal generator 505 on the basis of the application information) is discussed. At this time, if the motion vector of the partial region R_B is the same as that of the first sub-partition SP1, selection of either of the two candidates for the motion vector predictor results in the same motion vector predictor of the second sub-partition SP2 as the motion vector of the first sub-partition SP1. Therefore, it is preliminarily determined between the encoding device and the decoding device that if the two candidates for motion vector predictor are identical with each other, the predicted signal of the second sub-partition SP2 is to be generated by the motion vector resulting from addition of the differential motion vector and the motion vector predictor, which eliminates a need for transmitting the application information to indicate whether the differential motion vector is to be encoded, in addition to the indication information.

[0228] In cases where a target block is partitioned into three or more sub-partitions as shown in Fig. 33, the partition of the target block is meaningful if the first sub-partition SP1, second sub-partition SP2, and third sub-partition SP3 have the same motion vector and the fourth sub-partition SP4 only has a different motion vector. In such cases, therefore, the predicted signal of the second sub-partition SP2 and the predicted signal of the third sub-partition SP3 may be generated using the motion vector of the first sub-partition SP1, instead of motion vectors of partial region R_B and partial region R_E, respectively. However, for the fourth sub-partition SP4, if the motion vectors of the second sub-partition SP2 and the third sub-partition SP3 are the same,

two candidates for the motion vector predictor become identical with each other; therefore, by preliminarily determining a rule between the encoding device and the decoding device, it becomes unnecessary to transmit the indication information to indicate one motion vector predictor. Furthermore, if the first sub-partition SP1, second sub-partition SP2, and third sub-partition SP3 have the same motion vector and if the predicted signal of the fourth sub-partition SP4 is generated using the motion vector of the second sub-partition SP2 or the third sub-partition SP3, the four sub-partitions all will have the same motion vector; therefore, by preliminarily determining a rule between the encoding device and the decoding device, it also becomes unnecessary to transmit the application information to indicate whether the differential motion vector is to be encoded, in addition to the indication information.

[0229] A video decoding device according to an embodiment will be described below. Fig. 34 is a drawing schematically showing a configuration of the video decoding device according to an embodiment. The video decoding device 20 shown in Fig. 34 is a device that can generate a video sequence by decoding a compressed stream generated by the video encoding device 10.

[0230] As shown in Fig. 34, the video decoding device 20 is provided with data decoder 601, motion predictor 602, adder 603, inverse quantizer 604, inverse transformer 605, frame memory 606, predicted signal generator 607, and adder 608.

[0231] The data decoder 601 analyzes a compressed stream input via line L601. The data decoder 601 sequentially performs the processing

described below, for each block as a decoding target (hereinafter a target block).

[0232] The data decoder 601 decodes encoded data associated with the target block in the compressed stream to restore the quantized transform coefficients of the target block, and outputs the quantized transform coefficients via line L602. The data decoder 601 also decodes encoded data to restore the shape information of the target block, and outputs the shape information via line L603. On this occasion, the partition applicability information to indicate whether the target block needs to be partitioned is restored, and if the partition applicability information indicates no need for partition of the target block, the shape information does not have to be restored.

[0233] The data decoder 601 also decodes the encoded data to restore the indication information for each sub-partition in the target block, i.e., the information to indicate one of a plurality of candidates for motion vector predictor, and outputs the indication information via line L604. The data decoder 601 also decodes the encoded data to restore the differential motion vector of the target block, and outputs the differential motion vector via line L605. Furthermore, the data decoder 601 can adaptively determine the occurrence probability in decoding of encoded data on the occasion of restoring the indication information of the motion vector predictor, based on the shape information of the target block. A method for implementing it can be, for example, to set a higher occurrence probability for the indication information to indicate a motion vector of a partial region in contact with a processing target sub-partition, as a motion vector predictor.

[0234] The motion predictor 602 generates a motion vector predictor of a processing target sub-partition, based on the shape information fed via line L603 and motion vectors of partial regions prior in a processing order, fed via line L606, and based on the indication information fed via line L604, and outputs the motion vector predictor via line L607. By narrowing down candidates for the motion vector predictor to one by a predetermined method, it is also possible to omit the input of the indication information.

[0235] The adder 603 adds the motion vector predictor fed via line L607, to the differential motion vector fed via line L605, to generate a motion vector of a target block or a motion vector of a sub-partition in the target block, and outputs the motion vector via line L606.

[0236] The inverse quantizer 604 performs inverse quantization of the quantized transform coefficients fed via line L602, to generate inversely-quantized transform coefficients. The inverse quantizer 604 outputs the inversely-quantized transform coefficients thus generated, via line L608.

[0237] The inverse transformer 605 performs inverse orthogonal transformation of the inversely-quantized transform coefficients fed via line L608, to generate a reconstructed residual signal. The inverse transformer 605 outputs the reconstructed residual signal thus generated, via line L609.

[0238] If the reconstructed residual signal generated is not one subjected to quantization, the video decoding device 20 does not have to be provided with the inverse quantizer 604. Similarly, if the reconstructed residual signal generated is not one subjected to

transformation, the video decoding device 20 does not have to be provided with the inverse transformer 605.

[0239] The frame memory 606 stores previously-reconstructed image signals fed via line L610, i.e., frame image signals prior in the processing order to the processing target input image (which will be referred to hereinafter as reference frame image signals). Furthermore, the frame memory 606 outputs the reference frame image signals via line L611.

[0240] The predicted signal generator 607 generates a predicted signal of each sub-partition image in the target block, based on the motion vector fed via line L606 and the shape information fed via line L603, from an image signal in a predetermined range of the reference frame image signals fed via line L611. The predicted signal generator 607 outputs the predicted signal thus generated, via line L612. Although description is omitted in the present specification, the predicted signal may be generated by intra-picture prediction besides the inter-picture prediction.

[0241] The adder 608 adds the reconstructed residual signal fed via line L609, to the predicted signal of the target block fed via line L612, to generate a reconstructed image signal. The adder 608 outputs the reconstructed image signal via line L610.

[0242] Fig. 35 is a drawing showing a configuration of the motion predictor according to an embodiment. As shown in Fig. 35, the motion predictor 602 has a motion vector memory 6021, a motion reference candidate generator 6022, and a motion vector predictor generator 6023.

[0243] The motion vector memory 6021 stores motion vectors fed via line L606. The motion vectors stored in the motion vector memory 6021 are motion vectors of previously-processed partial regions prior in the processing order to the target block or the processing target sub-partition. The motion vector memory 6021 outputs the stored motion vectors via line L6021, for derivation of the motion vector predictor for the processing target sub-partition.

[0244] The motion reference candidate generator 6022 generates candidates for the motion vector predictor from the motion vectors fed via line L6021, by a below-described method, based on the shape information fed via line L603, and outputs them via line L6022.

[0245] The motion vector predictor generator 6023 determines a motion vector predictor, based on the indication information of the motion vector predictor fed via line L604, from the candidates for the motion vector predictor fed via line L6022, and outputs the determined motion vector predictor via line L607. If the number of candidates to be generated in the motion reference candidate generator is limited to one, the input of the indication information to specify the candidate to be selected can be omitted.

[0246] The below will describe the operation of the video decoding device 20 and a video decoding method according to an embodiment. Fig. 36 is a flowchart of a video decoding method according to an embodiment. In an embodiment, as shown in Fig. 36, in step S621, the data decoder 601 first decodes the encoded data in the compressed data about a target block, as described above, to restore the quantized transform coefficients, shape information, and differential motion vector

of the target block. In step S621, the partition applicability information and indication information may be restored. Furthermore, in step S621, the inverse quantizer 604 may generate the inversely-quantized transform coefficients from the quantized transform coefficients restored, and the inverse transformer 605 may generate a reconstructed residual signal from the inversely-quantized transform coefficients.

[0247] In step S622, the motion predictor 602 then determines the motion vector predictor of the processing target step S621, for each step S621 in the target block serving as a processing target. In subsequent step S623, the adder 603 adds the motion vector predictor of the processing target step S621 to the differential motion vector to generate a motion vector.

[0248] In step S624, the predicted signal generator 607 then generates the predicted signal from the reference frame image signals in the frame memory 606, using the motion vector of the target block. In subsequent step S625, the adder 608 adds the predicted signal of the target block to the reconstructed residual signal to generate a reconstructed image signal.

[0249] In step S626, the reconstructed image signal generated in step S625 is then stored as a previously-reconstructed image signal into the frame memory 606. In subsequent step S627, it is determined whether the processing is completed for all the blocks. If the processing is not completed for all the blocks, the processing from step S621 is continued using an unprocessed block as a target block. On the other hand, when the processing is completed for all the blocks, the processing is

terminated.

[0250] The operation of the motion predictor 602 will be described below in detail. Fig. 37 is a flowchart showing processing of the motion predictor according to an embodiment. The motion predictor 602 generates the motion vector predictor, according to the flowchart shown in Fig. 37.

[0251] In an embodiment, in step S615-1, the value of counter i is set to "0." It is assumed hereinafter that the processing for the first sub-partition is carried out with i=0 and the processing for the second sub-partition is carried out with i=1.

[0252] In next step S615-2, two candidates (left candidate and upper candidate) that can be the motion vector predictor of the processing target sub-partition are determined according to one of the methods described above using Figs. 29, 30, 31, 32, and 33, out of motion vectors of partial regions prior in the processing order to the processing target sub-partition.

[0253] In step S615-3, it is then determined whether the number NCand of candidates generated in step S615-2 is "0." When NCand is "0" (Yes), the processing proceeds to step S615-4. When NCand is not "0" (No), the processing proceeds to step S615-5.

[0254] In step S615-4, the motion vector predictor PMV is set to a zero vector and the processing proceeds to step S615-11. On this occasion, it is also possible to set a motion vector of a predetermined block or a motion vector of a partial region immediately prior in the processing order to the processing target sub-partition, instead of the zero vector, as the motion vector predictor PMV.

[0255] In step S615-5, it is determined whether the number NCand of candidates generated in step S615-2 is "1." When NCand is "1" (Yes), the processing proceeds to step S615-6. When NCand is not "1" (No), the processing proceeds to step S615-7.

[0256] In step S615-6, the candidate generated in step S615-2 is set as PMV. Then the processing proceeds to step S615-11.

[0257] In step S615-7, the information pmv_left_flag to indicate the PMV to be selected is acquired from the candidates generated in step S615-2. Then the processing proceeds to step S615-8.

[0258] In step S615-8, it is determined whether the value of pmv_left_flag is "1." When the value of pmv_left_flag is "1" (Yes), the processing proceeds to step S615-9. When the value of pmv_left_flag is not "1" (No), the processing proceeds to step S615-10.

[0259] Step S615-9 is to set a motion vector of a partial region on the left side of the processing target sub-partition as PMV. Then the processing proceeds to step S615-11.

[0260] Step S615-10 is to set a motion vector of a partial region on the left side of the processing target sub-partition as PMV. Then the processing proceeds to step S615-11.

[0261] Step S615-11 is to output the PMV thus set. Then the processing proceeds to step S615-12.

[0262] Next, step S615-12 is to add "1" to the value of counter i. Then the processing proceeds to step S615-13.

[0263] Next, step S615-13 is to determine whether the value of counter i is smaller than "2." When the value of counter i is smaller than "2" (Yes), the processing proceeds to step S615-2. On the other hand,

when the value of counter i is not smaller than 2 (No), the processing is terminated.

[0264] By limiting the number of candidates for the motion vector predictor to be generated, to one in step S615-2, the processes of steps S615-5, S615-6, S615-7, S615-8, S615-9, and S615-10 can be omitted. There are no restrictions on a method for this limitation, as described above as to the motion vector predictor generator 6023, but it is possible to use, for example, such a method as a method using an intermediate value of three candidates, a method using an average of two candidates, or a method of preliminarily determining a priority order for selection of one motion vector predictor out of a plurality of candidates for the motion vector predictor. In that case, when NCand is not "0" (No) in step S615-03, the processing proceeds to step S615-6.

[0265] The above-described method can be applied as a motion vector selection method in the case where the predicted signal of the processing target sub-partition is generated using previously-decoded motion vectors. Namely, the predicted signal of the processing target sub-partition may be generated using the motion vector predictor selected in step S615-2 in Fig. 37. In this case, there is no need for decoding of the differential motion vector, and therefore the motion vector predictor output from the motion predictor 602 is not output to the adder 603 but to the predicted signal generator 607.

[0266] Furthermore, the data decoder 601 may be configured to decode application information to specify whether the differential motion vector is to be decoded. In this modification, the motion predictor 602 may include a function to switch the output of the motion vector predictor

either to the adder 603 or to the predicted signal generator 607, based on the application information.

[0267] In this modification, it is unfavorable that the motion vectors of all the sub-partitions in the target block become identical with each other, because the partition of the target block becomes meaningless. In this modification, therefore, a motion vector of a sub-partition included in the target block and located prior in the processing order to the processing target sub-partition may be excluded from the candidates for the motion vector predictor, on the occasion of generating the candidates for the motion vector predictor of the processing target sub-partition in step S615-2 in Fig. 37. For example, in the case where the target block is partitioned into two sub-partitions and where the motion vector of the first sub-partition is first restored, the motion vector of the first sub-partitions is excluded from candidates for the motion vector predictor of the second sub-partition. If the motion vector of the first sub-partition is the same as that of the partial region U2, the motion vector of the partial region U2 does not have to be used in generation of the motion vector predictor of the second sub-partition.

[0268] In this modification, the occurrence probability in arithmetic decoding of the application information to indicate whether the differential motion vector is to be decoded can be adaptively determined according to the shape information. This method can be configured, for example, to set a higher probability of not encoding the differential motion vector, for the first sub-partition always being in contact with a previously-encoded region, than that for the second sub-partition possibly having no contact with any previously-decoded partial region.

Since the effect of this modification was already described using Figs. 34 and 35, the description thereof is omitted herein.

[0269] The below will describe a video encoding program for letting a computer operate as the video encoding device 10 and a video decoding program for letting a computer operate as the video decoding device 20.

[0270] Fig. 38 is a drawing showing a configuration of the video encoding program according to an embodiment. Fig. 39 is a drawing showing a configuration of the video decoding program according to an embodiment. Reference will be made below to Fig. 18 showing the hardware configuration of the computer according to an embodiment and Fig. 19 showing the perspective view of the computer according to an embodiment, as well as Figs. 38 and 39.

[0271] The video encoding program P10 shown in Fig. 38 can be provided as stored in the recording medium SM. The video decoding program P20 shown in Fig. 38 can also be provided as stored in the recording medium SM. Examples of recording media SM include recording media such as floppy disks, CD-ROMs, DVDs, or ROMs, or semiconductor memories, or the like.

[0272] As described above, the computer C10 may be provided with the reading device C12 such as a floppy disk drive unit, a CD-ROM drive unit, or a DVD drive unit, the working memory (RAM) C14 in which an operating system is resident, the memory C16 storing programs stored in the recording medium SM, the monitor device C18 such as a display, the mouse C20 and keyboard C22 as input devices, the communication device C24 for transmission and reception of data and others, and the CPU C26 to control execution of the programs.

[0273] When the recording medium SM is put into the reading device C12, the computer C10 becomes accessible to the video encoding program P10 stored in the recording medium SM, through the reading device C12, and becomes able to operate as the video encoding device 10, based on the program P10.

[0274] When the recording medium SM is put into the reading device C12, the computer C10 becomes accessible to the video decoding program P20 stored in the recording medium SM, through the reading device C12, and becomes able to operate as the video decoding device 20, based on the program P20.

[0275] As shown in Fig. 19, the video encoding program P10 and the video decoding program P20 may be those provided in the form of computer data signal CW superimposed on a carrier wave, through a network. In this case, the computer C10 can execute the program P10 or P20 after the video encoding program P10 or the video decoding program P20 received by the communication device C24 is stored into the memory C16.

[0276] As shown in Fig. 38, the video encoding program P10 includes block partition module M101, sub-partition generator module M102, storage module M103, motion detection module M104, predicted signal generation module M105, motion prediction module M106, subtraction module M107, residual signal generation module M108, transform module M109, quantization module M110, inverse quantization module M111, inverse transform module M112, addition module M113, and entropy encoding module M114.

[0277] In an embodiment, the block partition module M101,

sub-partition generator module M102, storage module M103, motion detection module M104, predicted signal generation module M105, motion prediction module M106, subtraction module M107, residual signal generation module M108, transform module M109, quantization module M110, inverse quantization module M111, inverse transform module M112, addition module M113, and entropy encoding module M114 cause the computer C10 to execute the same functions as the block partitioner 501, sub-partition generator 502, frame memory 503, motion detector 504, predicted signal generator 505, motion predictor 506, subtracter 507, residual signal generator 508, transformer 509, quantizer 510, inverse quantizer 511, inverse transformer 512, adder 513, and entropy encoder 514, respectively, in the video encoding device 10. Based on this video encoding program P10, the computer C10 becomes able to operate as the video encoding device 10.

[0278] As shown in Fig. 39, the video decoding program P20 includes data decoding module M201, motion prediction module M202, addition module M203, inverse quantization module M204, inverse transform module M205, storage module M206, predicted signal generation module M207, and addition module M208.

[0279] In an embodiment, the data decoding module M201, motion prediction module M202, addition module M203, inverse quantization module M204, inverse transform module M205, storage module M206, predicted signal generation module M207, and addition module M208 cause the computer C10 to execute the same functions as the data decoder 601, motion predictor 602, adder 603, inverse quantizer 604, inverse transformer 605, frame memory 606, predicted signal generator

607, and adder 608, respectively, in the video decoding device 20. Based on this video decoding program P20, the computer C10 becomes able to operate as the video decoding device 20.

[0280] A variety of embodiments were described above in detail. It is, however, noted that the present invention is by no means intended to be 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

[0281] 100 image predictive encoding device; 101 input terminal; 102 block partitioner; 103 predicted signal generator; 104 frame memory; 105 subtracter; 106 transformer; 107 quantizer; 108 inverse quantizer; 109 inverse transformer; 110 adder; 111 quantized transform coefficient encoder; 112 output terminal; 113 prediction block partitioning type selector; 114 motion information estimator; 115 prediction information memory; 116 prediction information encoder; 201 input terminal; 202 data analyzer; 203 inverse quantizer; 204 inverse transformer; 205 adder; 206 output terminal; 207 quantized transform coefficient decoder; 208 prediction information decoder; 10 video encoding device; 20 video decoding device; 501 block partitioner; 502 sub-partition generator; 503 frame memory; 504 motion detector; 505 predicted signal generator; 506 motion predictor; 507 subtracter; 508 residual signal generator; 509 transformer; 510 quantizer; 511 inverse quantizer; 512 inverse transformer; 513 adder; 514 entropy encoder; 601 data decoder; 602 motion predictor; 603 adder; 604 inverse quantizer; 605 inverse transformer; 606 frame memory; 607 predicted signal generator;

608 adder; 5061 motion vector memory; 5062 motion reference candidate generator; 5063 motion vector predictor generator; 6021 motion vector memory; 6022 motion reference candidate generator; 6023 motion vector predictor generator.

5

CLAIMS

1. An image predictive encoding device comprising:

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

prediction information estimation means adapted to

sub-partition a target region serving as an encoding target partitioned by the region partition means into a first prediction region and a second prediction region,

determine a prediction block partitioning type indicating a number and region shapes of the prediction regions suitable for the target region,

predict first and second motion information respectively for acquiring a signal highly correlated with the first prediction region and a signal highly correlated with the second prediction region based on previously-reconstructed signals, and

obtain prediction information including:

the prediction block partitioning type,

the first motion information,

the second motion information,

a first merge identification information indicating whether or not to utilize decoded motion information associated with neighboring regions neighboring the first prediction region for generating a predicted signal of the first prediction region, and

a second merge identification information indicating whether or not to utilize decoded motion information except the first motion information from among the decoded motion information associated with neighboring regions neighboring the second prediction region for generating a predicted signal of the second prediction region;

prediction information encoding means adapted to encode the prediction information associated with the target region;

predicted signal generation means adapted to generate a predicted signal of each of the first and second prediction regions based on the first motion information and the second motion information;

residual signal generation means adapted to generate a residual signal based on a pixel signal and the predicted signal of each of the first and second prediction regions;

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

residual signal restoration means adapted to decode the encoded residual signal to generate a reconstructed residual signal; and

recording means adapted to generate a restored pixel signal of the target region based on the predicted signal and the reconstructed residual signal and to store the restored pixel signal as the previously-reconstructed signal.

2. An image predictive encoding method executed by an image predictive encoding device, comprising:

a region partition step of partitioning an input image into a plurality of regions;

a prediction information estimation step of

sub-partitioning a target region serving as an encoding target partitioned by the region partition step into a first prediction region and a second prediction region,

determining a prediction block partitioning type indicating a number and region shapes of the prediction regions suitable for the target region,

predicting first and second motion information respectively for acquiring a signal highly correlated with the first prediction region and a signal highly correlated with the second prediction region based on previously-reconstructed signals, and

obtaining prediction information including:

the prediction block partitioning type,

the first motion information,

the second motion information,

a first merge identification information indicating whether or not to utilize decoded motion information associated with neighboring regions neighboring the first prediction region for generating a predicted signal of the first prediction region, and

a second merge identification information indicating whether or not to utilize decoded motion information except the first motion information from among the decoded motion information associated with neighboring regions neighboring the second prediction region for generating a predicted signal of the second prediction region;

a prediction information encoding step of encoding the prediction information associated with the target region;

a predicted signal generation step of generating a predicted signal of each of the first and second prediction regions based on the first motion information and the second motion information;

a residual signal generation step of generating a residual signal based on a pixel signal and the predicted signal of each of the first and second prediction regions;

a residual signal encoding step of encoding the residual signal generated by the residual signal generation step;

a residual signal restoration step of decoding the encoded residual signal to generate a reconstructed residual signal; and

a recording step of generating a restored pixel signal of the target region based on the predicted signal and the reconstructed residual signal and of storing the restored pixel signal as the previously-reconstructed signal.

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

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

prediction information estimation means adapted to

sub-partition a target region serving as an encoding target partitioned by the region partition means into a first prediction region and a second prediction region,

determine a prediction block partitioning type indicating a number and region shapes of the prediction regions suitable for the target region,

predict first and second motion information respectively for acquiring a signal highly correlated with the first prediction

region and a signal highly correlated with the second prediction region based on previously-reconstructed signals, and

obtain prediction information including:

the prediction block partitioning type,

the first motion information,

the second motion information,

a first merge identification information indicating whether or not to utilize decoded motion information associated with neighboring regions neighboring the first prediction region for generating a predicted signal of the first prediction region, and

a second merge identification information indicating whether or not to utilize decoded motion information except the first motion information from among the decoded motion information associated with neighboring regions neighboring the second prediction region for generating a predicted signal of the second prediction region;

prediction information encoding means adapted to encode the prediction information associated with the target region;

predicted signal generation means adapted to generate a predicted signal of each of the first and second prediction regions based on the first motion information and the second motion information;

residual signal generation means adapted to generate a residual signal based on a pixel signal and the predicted signal of each of the first and second prediction regions;

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

residual signal restoration means adapted to decode the encoded residual signal to generate a reconstructed residual signal; and

recording means adapted to generate a restored pixel signal of the target region based on the predicted signal and the reconstructed residual signal and to store the restored pixel signal as the previously-reconstructed signal.

4. An image predictive decoding device comprising:

data analysis means adapted to extract encoded data of prediction

information to indicate a prediction method to be used in prediction of a signal of a target region serving as a decoding target and encoded data of a residual signal, from compressed image data obtained by partitioning an image into a plurality of regions and encoding image data of the regions;

prediction information decoding means adapted to restore motion information based on the encoded data of prediction information,

wherein the prediction information decoding means restores a prediction block partitioning type indicating a number of prediction regions obtained by sub-partitioning the target region, based on the encoded data of prediction information,

wherein, when the prediction block partitioning type indicates that the target region includes a first prediction region and a second prediction region, the prediction information decoding means further decodes the encoded data of prediction information to restore a first merge identification information indicating whether or not to utilize decoded motion information associated with neighboring regions neighboring the first prediction region for generating a predicted signal of the first prediction region,

wherein, when the first merge identification information indicates not to utilize the decoded motion information, the prediction information decoding means further decodes the encoded data of prediction information to restore a first motion information utilized for generating the predicted signal of the first prediction region,

wherein, when the first merge identification information indicates to utilize the decoded motion information, the prediction information decoding means further decodes the encoded data of prediction information to restore a first selection information identifying the first motion information utilized for generating the predicted signal of the first prediction region based on the decoded motion information associated with neighboring regions neighboring the first prediction region and to restore the first motion information based on the first selection information,

wherein, the prediction information decoding means further

decodes the encoded data of prediction information to restore a second merge identification information indicating whether or not to utilize decoded motion information associated with neighboring regions neighboring the second prediction region for generating a predicted signal of the second prediction region,

wherein, when the second merge identification information indicates not to utilize the decoded motion information, the prediction information decoding means further decodes the encoded data of prediction information to restore a second motion information utilized for generating the predicted signal of the second prediction region, and

wherein, when the second merge identification information indicates to utilize the decoded motion information, the prediction information decoding means further decodes the encoded data of prediction information to restore a second selection information identifying the second motion information utilized for generating the predicted signal of the second prediction region based on candidate for the decoded motion information except the first motion information from among the decoded motion information associated with neighboring regions neighboring the second prediction region, and to restore the second motion information based on the second selection information;

storing means adapted to store motion information included in restored prediction information;

predicted signal generation means adapted to generate a predicted signal of each of the first and second prediction regions in the target region based on the restored first motion information and the restored second motion information;

residual signal restoration means adapted to restore a reconstructed residual signal of the target region based on the encoded data of the residual signal; and

recording means adapted to generate a restored pixel signal of the target region based on the predicted signal and the reconstructed residual signal and to store the restored pixel signal as the previously-reconstructed signal.

5. An image predictive decoding method executed by an image predictive decoding device, comprising:

a data analysis step of extracting encoded data of prediction information to indicate a prediction method to be used in prediction of a signal of a target region serving as a decoding target and encoded data of a residual signal, from compressed image data obtained by partitioning an image into a plurality of regions and encoding image data of the regions;

a prediction information decoding step of restoring motion information based on the encoded data of prediction information,

wherein the image predictive decoding device restores a prediction block partitioning type indicating a number of prediction regions obtained by sub-partitioning the target region, based on the encoded data of prediction information,

wherein, when the prediction block partitioning type indicates that the target region includes a first prediction region and a second prediction region, the image predictive decoding device further decodes the encoded data of prediction information to restore a first merge identification information indicating whether or not to utilize decoded motion information associated with neighboring regions neighboring the first prediction region for generating a predicted signal of the first prediction region,

wherein, when the first merge identification information indicates not to utilize the decoded motion information, the image predictive decoding device further decodes the encoded data of prediction information to restore a first motion information utilized for generating the predicted signal of the first prediction region,

wherein, when the first merge identification information indicates to utilize the decoded motion information, the image predictive decoding device further decodes the encoded data of prediction information to restore a first selection information identifying the first motion information utilized for generating the predicted signal of the first prediction region based on the decoded motion information associated with neighboring regions neighboring the first prediction region and to restore the first

motion information based on the first selection information,

wherein, the image predictive decoding device further decodes the encoded data of prediction information to restore a second merge identification information indicating whether or not to utilize decoded motion information associated with neighboring regions neighboring the second prediction region for generating a predicted signal of the second prediction region,

wherein, when the second merge identification information indicates not to utilize the decoded motion information, the image predictive decoding device further decodes the encoded data of prediction information to restore a second motion information utilized for generating the predicted signal of the second prediction region, and

wherein, when the second merge identification information indicates to utilize the decoded motion information, the image predictive decoding device further decodes the encoded data of prediction information to restore a second selection information identifying the second motion information utilized for generating the predicted signal of the second prediction region based on candidate for the decoded motion information except the first motion information from among the decoded motion information associated with neighboring regions neighboring the second prediction region, and to restore the second motion information based on the second selection information;

a storing step of storing motion information included in restored prediction information;

a predicted signal generation step of generating a predicted signal of each of the first and second prediction regions in the target region based on the restored first motion information and the restored second motion information;

a residual signal restoration step of restoring a reconstructed residual signal of the target region based on the encoded data of the residual signal; and

a recording step of generating a restored pixel signal of the target region based on the predicted signal and the reconstructed residual signal

and to store the restored pixel signal as the previously-reconstructed signal.

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

data analysis means adapted to extract encoded data of prediction information to indicate a prediction method to be used in prediction of a signal of a target region serving as a decoding target and encoded data of a residual signal, from compressed image data obtained by partitioning an image into a plurality of regions and encoding image data of the regions;

prediction information decoding means adapted to restore motion information based on the encoded data of prediction information,

wherein the prediction information decoding means restores a prediction block partitioning type indicating a number of prediction regions obtained by sub-partitioning the target region, based on the encoded data of prediction information,

wherein, when the prediction block partitioning type indicates that the target region includes a first prediction region and a second prediction region, the prediction information decoding means further decodes the encoded data of prediction information to restore a first merge identification information indicating whether or not to utilize decoded motion information associated with neighboring regions neighboring the first prediction region for generating a predicted signal of the first prediction region,

wherein, when the first merge identification information indicates not to utilize the decoded motion information, the prediction information decoding means further decodes the encoded data of prediction information to restore a first motion information utilized for generating the predicted signal of the first prediction region,

wherein, when the first merge identification information indicates to utilize the decoded motion information, the prediction information decoding means further decodes the encoded data of prediction information to restore a first selection information identifying the first motion information utilized for generating the predicted signal of the first prediction region

based on the decoded motion information associated with neighboring regions neighboring the first prediction region and to restore the first motion information based on the first selection information,

wherein, the prediction information decoding means further decodes the encoded data of prediction information to restore a second merge identification information indicating whether or not to utilize decoded motion information associated with neighboring regions neighboring the second prediction region for generating a predicted signal of the second prediction region,

wherein, when the second merge identification information indicates not to utilize the decoded motion information, the prediction information decoding means further decodes the encoded data of prediction information to restore a second motion information utilized for generating the predicted signal of the second prediction region, and

wherein, when the second merge identification information indicates to utilize the decoded motion information, the prediction information decoding means further decodes the encoded data of prediction information to restore a second selection information identifying the second motion information utilized for generating the predicted signal of the second prediction region based on candidate for the decoded motion information except the first motion information from among the decoded motion information associated with neighboring regions neighboring the second prediction region, and to restore the second motion information based on the second selection information;

storing means adapted to store motion information included in restored prediction information;

predicted signal generation means adapted to generate a predicted signal of each of the first and second prediction regions in the target region based on the restored first motion information and the restored second motion information;

residual signal restoration means adapted to restore a reconstructed residual signal of the target region based on the encoded data of the

residual signal; and

recording means adapted to generate a restored pixel signal of the target region based on the predicted signal and the reconstructed residual signal and to store the restored pixel signal as the previously-reconstructed signal.