The motion prediction mode generation section 51 supplies information indicating the inter prediction mode and a reference image index (Ref index) to the AMVP mode vector prediction section 53. The motion prediction mode generation section 51 supplies the merge mode or the skip mode (M/S mode) to the automatic reference index generation section 52.

[0089]

The automatic reference index generation section 52 automatically generates a reference image index, and supplies the generated reference image index (Ref index) to the M/S mode vector prediction section 54 along with the merge mode or the skip mode from the motion prediction mode generation section 51.

[0090]

The AMVP mode vector prediction section 53 reads one or more pictures encoded before a target picture and locally decoded from the DPB 32-1 or 32-N according to the prediction mode from the motion prediction mode generation section 51 and the reference image index as the candidate pictures.

[0091]

The AMVP mode vector prediction section 53 detects an MV representing motion serving as a shift between a target block and a corresponding block corresponding to the target block in a candidate picture by motion detection using the target block of a target picture from the picture reordering buffer 22 and the candidate picture. The AMVP mode vector prediction section 53 generates a predicted image by performing motion compensation for compensating for a shift of

pictures of different times of the same view from the MV memory 36-1 and designates the read MVs as candidates for the predicted vector. The AMVP mode

vector prediction section 53 reads MVs of reference blocks in different views of the same time from the MV memory 36-N in which the MVs of the different views are stored, and designates the read MVs as the candidates for the predicted vector.

[0093]

The AMVP mode vector prediction section 53 obtains an encoding cost required to encode a target block using a predicted image for every candidate picture for use in generation of a predicted image, every candidate for an MV, every candidate for a predicted vector, or every inter prediction mode with a variable block size based on an original image from the image reordering buffer 22. The AMVP mode vector prediction section 53 supplies an optimal encoding cost among obtained encoding costs as a mode cost to the mode determination section 55. At this time, the AMVP mode vector prediction section 53 designates a candidate for the predicted vector at the time of the optimal encoding cost as a predicted vector, obtains a difference from the MV, and encodes the MV difference Mvd and an index (Mv index) of the predicted vector as MV information.

[0094]

The M/S mode vector prediction section 54 reads one or more pictures encoded before a target picture and locally decoded from the DPB 32-1 or 32-N according to the mode and the reference image index from the automatic reference index generation section 52 as the candidate pictures.

[0095]

In addition, the M/S mode vector prediction section 54 designates a

read MVs as the candidates for the MV. The M/S mode vector prediction section 54 generates a candidate picture using the candidates for the MV.

[0096]

The M/S mode vector prediction section 54 obtains an encoding cost required to encode a target block using a predicted image for every candidate picture for use in generation of a predicted image, every candidate for an MV, or every M/S mode based on an original image from the image reordering buffer 22. The M/S mode vector prediction section 54 supplies an optimal encoding cost among the obtained encoding costs as a mode cost to the mode determination section 55. In addition, the M/S mode vector prediction section 54 encodes a merge index indicating an MV as MV information.

[0097]

The mode determination section 55 refers to encoding costs from the AMVP mode vector prediction section 53 and the M/S mode vector prediction section 54 and determines the inter prediction mode or the inter-view prediction mode having a lowest encoding cost as an optimum prediction mode which is an optimal motion prediction mode. The mode determination section 55 returns an optimum prediction mode determination result to the AMVP mode vector prediction section 53 and the M/S mode vector prediction section 54.

[0098]

The AMVP mode vector prediction section 53 supplies the predicted image (pred. image) and the encoding cost obtained in the optimum prediction mode to the predicted image selection section 35 based on the determination result from the mode determination section 55. The AMVP mode vector prediction section 53 supplies

determination section 55. In addition, the M/S mode vector prediction section 54 supplies the prediction mode (M/S mode) determined as the optimum prediction

mode and the encoded MV information to the lossless encoding section 26. At this time, the information of the MV of the optimum encoding cost is temporarily saved in (overwritten on) the spatial MV memory of FIG. 6 as will be described later.

[0100]

[Configuration of AMVP mode vector prediction section]

FIG. 6 is a block diagram illustrating a configuration example of the AMVP mode vector prediction section of FIG. 5.

[0101]

In the example of FIG. 6, the AMVP mode vector prediction section 53 is configured to include a vector search section 61, a predicted image generation section 62, a vector cost determination section 63, a spatial MV memory 64, predicted vector generation sections 65 and 66, a switch 67, a subtraction section 68, and a POC transform section 69.

[0102]

The reference image index from the motion prediction mode generation section 51 is supplied to the vector search section 61, the POC transform section 69, and the lossless encoding section 26. In addition, the prediction mode from the motion prediction mode generation section 51 is also supplied to the vector search section 61.

[0103]

The vector search section 61 reads one or more pictures encoded before a target picture and locally decoded from the DPB 32-1 or 32-N according to the

determination section 63.

[0104]

The predicted image generation section 62 generates a predicted image by performing motion compensation for compensating for a shift of a motion amount of the candidate picture from the DPB 32-1 or 32-N according to the MV of the target block from the vector search section 61. The generated predicted image is supplied to the predicted image selection section 35 and the vector cost determination section 63.

[0105]

The vector cost determination section 63 obtains an encoding cost using an original image from the picture reordering buffer 22, an MV from the vector search section 61, a predicted image from the predicted image generation section 62, and predicted vectors and their MV indices from the predicted vector generation sections 65 and 66. Then, the vector cost determination section 63 determines a lowest encoding cost and supplies the lowest encoding cost (best cost) and its prediction mode to the mode determination section 55. The vector cost determination section 63 temporarily saves the MV of the lowest encoding cost in the spatial MV memory 64.

[0106]

The spatial MV memory 64 stores the MV of the lowest encoding cost as the candidate for use in generation of the predicted vector to be performed at a later time. In the spatial MV memory 64, the MV is stored in each unit (PU) of a block in which the MV is obtained. Also, when the encoding cost of the M/S mode is optimum, the MV of the spatial MV memory 64 is overwritten on the MV of the case

vector by reading MVs of spatially neighboring blocks adjacent thereto in the same picture. The predicted vector generation section 65 supplies the generated spatially

predicted vector to the vector cost determination section 63 and the subtraction section 68 via the switch 67 along with the MV index indicating the predicted vector.

[0109]

The predicted vector generation section 66 generates a predicted vector by temporal moving vector prediction (TMVP). That is, the predicted vector generation section 66 generates the predicted vector by reading MVs of temporally neighboring blocks corresponding or adjacent thereto in pictures of different times of the same view from the MV memory 36-1. Also, at this time, based on POC information from the POC transform section 69, scaling is performed when the reference POC (Ref POC) of the target block is different from a reference POC (Ref POC) of a temporally neighboring block. That is, the scaled MV serves as the predicted vector. The predicted vector generation section 66 supplies the generated temporally predicted vector to the vector cost determination section 63 and the subtraction section 68 via the switch 67 along with an MV index indicating the predicted vector.

[0110]

In addition, the predicted vector generation section 66 generates a predicted vector by inter-view prediction (IVMP). The predicted vector generation section 66 finds a disparity vector from MVs of neighboring blocks adjacent to the target block from the spatial MV memory 64 and obtains reference blocks in different views of the same time based on the found disparity vector. Then, the predicted vector generation section 66 generates a predicted vector by reading MVs of reference

cost determination section 63 and the subtraction section 68 via the switch 67 along with an MV index indicating the predicted vector.

[0112]

The switch 67 selects the predicted vector from the predicted vector generation section 65 or the predicted vector from the predicted vector generation section 66, and supplies the selected predicted vector and its MV index to the vector cost determination section 63 and the subtraction section 68.

[0113]

The subtraction section 68 encodes a difference MVd between the MV (best MV) having the lowest cost from the spatial MV memory 64 and the predicted vector from the switch 67 as MV information along with an MV index representing an index of the predicted vector. The subtraction section 68 supplies the encoded MV information to the lossless encoding section 26.

[0114]

The POC transform section 69 transforms a reference image index (Ref index) of the target block from the motion prediction mode generation section 51 into a POC and supplies POC information indicating the POC obtained by the transform to the predicted vector generation section 66.

[0115]

[Configuration example of non-spatially predicted vector generation section]

FIG. 7 is a block diagram illustrating the configuration example of the non-spatially predicted vector generation section of FIG. 6.

[0116]

index to the intra-view reference vector generation section 82. The predicted vector index generation section 81 generates a predicted vector index (MV index) of the

IVMP and supplies the generated predicted vector index to the inter-view reference vector generation section 83.

[0118]

The intra-view reference vector generation section 82 generates a predicted vector by the TMVP. That is, the intra-view reference vector generation section 82 generates the predicted vector by reading MVs of corresponding blocks associated by MVs in pictures of different times of the same view from the MV memory 36-1.

[0119]

Also, at this time, based on POC information from the POC transform section 69, scaling is performed on an MV of a corresponding block when the reference POC (Ref POC) of the target block is different from a reference POC (Ref POC) of the corresponding block. That is, the scaled MV serves as the predicted vector. The intra-view reference vector generation section 82 supplies the generated temporally predicted vector (PMV) to the vector cost determination section 63 and the subtraction section 68 via the switch 67 along with an MV index indicating the predicted vector.

[0120]

The inter-view reference vector generation section 83 generates a predicted vector by IVMP. The inter-view reference vector generation section 83 finds a disparity vector from MVs of neighboring blocks adjacent to the target block, and obtains a reference block in different views of the same time based on the found disparity vector. Then, the inter-view reference vector generation section 83 section 69. That is, the scaled MV serves as the predicted vector. The inter-view reference vector generation section 83 supplies the generated inter-view predicted

vector to the vector cost determination section 63 and the subtraction section 68 via the switch 67 along with an MV index indicating the predicted vector.

[0122]

[Operation of encoder]

Next, an encoding process of the encoder 11-1 of FIG. 4 will be described with reference to a flowchart of FIG. 8. Also, the encoders 11-N and 11-M for performing processes of encoding images of other views perform similar encoding processes.

[0123]

A picture of the color image of the non-base view that is an image (moving image) to be encoded is sequentially supplied to the A/D conversion section 21 in the display order. In step S11, when the picture is an analog signal, the A/D conversion section 21 converts the analog signal according to A/D conversion and supplies an A/D conversion result to the picture reordering buffer 22.

[0124]

The picture reordering buffer 22 temporarily stores a picture from the A/D conversion section 21 and reads a picture according to a structure of a GOP indicated by the supplied encoding order, so that a process of reordering a picture sequence from the display order to the encoding order (decoding order) is performed. The picture read from the picture reordering buffer 22 is supplied to the calculation section 23, the intra-picture prediction section 33, and the motion prediction/compensation section 34.

[0126]

The intra-picture prediction section 33 obtains an encoding cost required to

encode the target block using the predicted image, that is, an encoding cost required to encode a residual or the like for the predicted image in the target block, and supplies the obtained encoding cost to the predicted image selection section 35 along with the predicted image.

[0127]

In step S13, the motion prediction/compensation section 34 performs motion prediction and compensation. Also, the motion prediction and compensation process will be described in detail with reference to FIG. 9.

[0128]

In step S13, motion prediction, compensation, and predicted vector generation in all inter prediction modes, MV generation in the M/S mode, and the like are performed and a predicted image is generated in all the inter prediction modes (including the M/S mode). Then, the encoding cost required to encode the target block (PU) using the predicted image is obtained for every candidate picture for use in generation of a predicted image, every candidate for an MV, every candidate for a predicted vector, or every inter prediction mode (including the inter-view prediction mode) with a variable block size, or every M/S mode, an optimum inter prediction mode is determined, and the encoding cost is supplied to the predicted image selection section 35 along with the predicted image.

[0129]

Also, at this time, the intra-picture prediction section 33 supplies information about intra prediction as header information to the lossless encoding

supplies the selected predicted image to the calculation sections 23 and 30.

[0131]

In step S15, the motion prediction/compensation section 34 (the vector cost determination section 63 of FIG. 6) temporarily saves the MV of the optimum inter prediction mode when motion prediction (other than intra-picture prediction) is selected in step S14 in the spatial MV memory 64 of FIG. 6. That is, although the MV of the case of the AMVP mode is saved in the spatial MV memory 64 even in the process of step S13, the MV of the spatial MV memory 64 of FIG. 6 becomes the MV of the M/S mode, for example, when a mode of an optimal encoding cost is the M/S mode in step S15.

[0132]

In step S16, the calculation section 23 computes a difference between the original image from the picture reordering buffer 22 and the predicted image from the predicted image selection section 35, and supplies a computation result to the orthogonal transform section 24. That is, the calculation section 23 performs prediction encoding after calculating a subtraction value by subtracting a pixel value of the predicted image supplied from the predicted image selection section 35 from the pixel value of the target block if necessary, and supplies a prediction encoding result to the orthogonal transform section 24.

[0133]

In step S17, the orthogonal transform section 24 performs an orthogonal transform such as a discrete cosine transform or a Karhunen-Loeve transform on (a residual obtained by subtracting the predicted image or a pixel value of) a target block from the calculation section 23 using a transform unit (TU) as a unit, and

[0135]

In step S19, the inverse quantization section 28 inversely quantizes the

quantization value from the quantization section 25 into a transform coefficient and supplies the transform coefficient to the inverse orthogonal transform section 29.

[0136]

In step S20, the inverse orthogonal transform section 29 performs an inverse orthogonal transform on the transform coefficient from the inverse quantization section 28 and supplies the transformed transform coefficient to the calculation section 30.

[0137]

In step S21, the lossless encoding section 26 performs lossless encoding on a residual coefficient that is the quantization value from the quantization section 25 and supplies encoded data obtained as a result thereof to the accumulation buffer 27. Also, the lossless encoding section 26 encodes header information such as prediction mode information or MV information from the intra-picture prediction section 33 or the motion prediction/compensation section 34 and includes the encoded header information in the header of the encoded data.

[0138]

In step S22, the calculation section 30 obtains a decoded image in which a target block is decoded (locally decoded) by adding the pixel value of the predicted image supplied from the predicted image selection section 35 to the data supplied from the inverse orthogonal transform section 29 if necessary, and supplies the obtained decoded image to the in-loop filter 31.

[0139]

section 30.

[0141]

In step S25, the in-loop filter 31 saves the decoded image after the filtering in the DPB 32-1.

[0142]

In step S26, the motion prediction/compensation section 34 compresses an MV saved in step S15. That is, for example, although one MV is saved for every (4×4) block in the spatial MV memory 64, the MV is compressed so that one MV is saved for every (16×16) block. For example, an MV of an upper-left block is selected in the (16×16) block.

[0143]

Then, the motion prediction/compensation section 34 saves the compressed MV in the MV memory 36-1 in step S27.

[0144]

In step S28, the encoder 11-M of the other view encodes a picture of the other view. Also, the encoding process is basically similar to the encoding process of FIG. 8.

[0145]

As described above, the encoding process is performed.

[0146]

[Example of motion prediction/compensation process]

Next, the motion prediction/compensation process of step S13 of FIG. 8 will be described with reference to the flowchart of FIG. 9.

[0147]

in step S42, the motion prediction mode generation section 51 supplies the inter prediction mode (inter mode) and a reference image index (Ref index) to the AMVP

mode vector prediction section 53. Then, the process proceeds to step S43.

[0149]

In step S43, the AMVP mode vector prediction section 53 performs vector prediction of the AMVP mode. Details of the vector prediction of the AMVP mode will be described later with reference to FIG. 10.

[0150]

In the process of step S43, an MV in the inter prediction mode is found, a predicted image is generated, and a residual image is generated, each of spatially and non-spatially predicted vectors is generated. In particular, if the Ref POC of the current PU is different from the Ref POC of the reference PU in the different view when the non-spatially predicted vector is generated, the MV of the reference PU is scaled and the scaled MV serves as a candidate for the predicted vector of the current PU. Then, a difference between the candidate for the predicted vector and the MV is calculated, so that the predicted vector of a lowest cost is selected. The lowest cost of the selected predicted vector is supplied to the mode determination section 55. In addition, a difference between the predicted vector of the selected lowest cost and the MV and an index of the predicted vector are encoded as MV information.

[0151]

On the other hand, when it is determined that the mode is not the inter prediction mode in step S42, the motion prediction mode generation section 51 supplies the merge mode or the skip mode (M/S mode) to the automatic reference index generation section 52. Then, the process proceeds to step S44.

In step S45, the M/S mode vector prediction section 54 performs vector prediction of the merge mode or the skip mode.

[0154]

That is, the M/S mode vector prediction section 54 reads one or more pictures encoded before a target picture and locally decoded from the DPB 32-1 or 32-N according to the mode and the reference image index from the automatic reference index generation section 52 as the candidate pictures.

[0155]

In addition, the M/S mode vector prediction section 54 designates a candidate for the MV using spatially neighboring blocks adjacent thereto in the same picture. The M/S mode vector prediction section 54 reads MVs of temporally neighboring blocks corresponding or adjacent thereto in pictures of different times of the same view from the MV memory 36-1 and designates the read MVs as candidates for the MV. The M/S mode vector prediction section 54 reads MVs of reference blocks in different views of the same time from the MV memory 36-N in which the MVs of the different views are stored, and designates the read MVs as the candidates for the predicted vector. The M/S mode vector prediction section 54 generates a candidate picture using the MV candidates.

[0156]

The M/S mode vector prediction section 54 obtains an encoding cost required to encode a target block using a predicted image for every candidate picture for use in generation of a predicted image, every candidate for an MV, or every M/S mode based on an original image from the image reordering buffer 22. The M/S mode vector prediction section 54 supplies an optimal encoding cost among the prediction mode having a lowest encoding cost as an optimum prediction mode which is an optimum motion prediction mode. The mode determination section 55

returns an optimum inter prediction mode determination result to the AMVP mode vector prediction section 53 and the M/S mode vector prediction section 54.

[0158]

In step S47, the AMVP mode vector prediction section 53 or the M/S mode vector prediction section 54 selects encoded motion information of a mode having a low encoding cost based on a determination result from the mode determination section 55 and supplies the selected motion information to the lossless encoding section 26.

[0159]

[Example of vector prediction process of AMVP mode]

Next, the vector prediction process of the AMVP mode of step S43 of FIG. 9 will be described with reference to a flowchart of FIG. 10.

[0160]

The prediction mode from the motion prediction mode generation section 51 is supplied to the vector search section 61 along with a reference image index.

[0161]

In step S61, the vector search section 61 performs a vector search according to the prediction mode and the reference image index from the motion prediction mode generation section 51.

[0162]

That is, the vector search section 61 reads one or more pictures encoded before a target picture and locally decoded from the DPB 32-1 or 32-N according to

determination section 63.

[0163]

In step S62, the predicted image generation section 62 generates a predicted image according to the MV of the target block from the vector search section 61.

[0164]

That is, the predicted image generation section 62 generates the predicted image by performing motion compensation for compensating for a shift of a motion amount of the candidate picture from the DPB 32-1 or 32-N according to the MV of the target block from the vector search section 61. The generated predicted image is supplied to the predicted image selection section 35 and the vector cost determination section 63.

[0165]

In step S63, the vector determination section 63 generates a residual image using the original image from the picture reordering buffer 22, the MV from the vector search section 61, and the predicted image from the predicted image generation section 62. The generated residual image is used to compute the encoding cost in step S67 as will be described later.

[0166]

In step S64, the predicted vector generation section 65 generates a spatially predicted vector. That is, the predicted vector generation section 65 generates the spatially predicted vector by reading MVs of spatially neighboring blocks adjacent thereto in the same picture from the spatial MV memory 64. The predicted vector generation section 65 supplies the generated spatially predicted vector to the vector cost determination section 63 and the subtraction section 68 via the switch 67 along

[0168]

In the process of step S65, a process of generating the predicted vector by

the TMVP is performed and a process of generating the predicted vector by the IVMP is performed. Also, at the time of the process of generating the predicted vector by the IVMP, a disparity vector is found from the MVs of neighboring blocks adjacent to the target block and reference blocks in different views of the same time are obtained based on the found disparity vector. Then, the MVs of the reference blocks in the different views of the same time are read from the MV memory 36-N in which the MVs of the different views are stored, and scaling is performed when the reference POC (Ref POC) of the target block is different from a reference POC (Ref POC) of the reference block. Also, the POC information is transformed from the Ref index in the POC transform section 69 and the transformed POC information is supplied.

[0169]

The predicted vector by the TMVP and the predicted vector of the IVMP generated in the process of step S65 are supplied to the vector cost determination section 63 and the subtraction section 68 via the switch 67 along with an MV index indicating the predicted vector.

[0170]

In step S66, the vector cost determination section 63 computes residuals between the MV of the target block and the predicted vectors of the target block supplied from the predicted vector generation sections 65 and 66.

[0171]

In step S67, the vector cost determination section 63 obtains encoding costs

[0173]

In step S68, a difference MVd between the MV (best MV) having the lowest

cost from the spatial MV memory 64 and the predicted vector from the switch 67 corresponding to the best MV is encoded as MV information along with an MV index representing an index of the predicted vector.

[0174]

[Example of process of generating non-spatially predicted vector]

Next, the process of generating the non-spatially predicted vector in step S65 of FIG. 10 will be described with reference to the flowchart of FIG. 11.

[0175]

The predicted vector index generation section 81 generates a predicted vector index (MV index) of the TMVP and supplies the generated predicted vector index to the intra-view reference vector generation section 82. The predicted vector index generation section 81 generates a predicted vector index (MV index) of the IVMP and supplies the generated predicted vector index to the inter-view reference vector generation section 83.

[0176]

In step S81, the intra-view reference vector generation section 82 generates a predicted vector by the TMVP.

[0177]

That is, the intra-view reference vector generation section 82 generates the predicted vector by reading MVs of temporally neighboring blocks corresponding or adjacent thereto in pictures of different times of the same view from the MV memory 36-1. The intra-view reference vector generation section 82 supplies the generated

That is, in step S82, the inter-view reference vector generation section 83 finds a disparity vector from MVs of neighboring blocks adjacent to a target block

(PU) from the spatial MV memory 64, and calculates a disparity based on the found disparity vector.

[0180]

In step S83, the inter-view reference vector generation section 83 selects a PU of a position shifted by the disparity obtained in step S82 as the reference PU in a different view.

[0181]

In step S84, the inter-view reference vector generation section 83 generates a predicted vector from an MV of a selected reference PU by reading the MV of the selected reference PU from the MV memory 36-N in which MVs of the different views are stored. This predicted vector generation process will be described later with reference to FIGS. 12 and 13.

[0182]

In the process of step S84, the generated inter-view predicted vector of the IVMP is supplied to the vector cost determination section 63 and the subtraction section 68 via the switch 67 along with an MV index indicating the predicted vector.

[0183]

[Example of predicted vector generation process]

Next, the predicted vector generation process of step S84 of FIG. 11 will be described with reference to a flowchart of FIG. 12. Also, in the example of FIG. 12, a predicted vector generation process in a direction L0 is illustrated.

[0184]

S102. In step S102, the inter-view reference vector generation section 83 determines whether the POCcurr_l0 that is a Ref POC of the target PU is equal to

POCbase_l0 that is a Ref POC of the reference PU.

[0186]

When it is determined that POCcurr_l0 is equal to POCbase_l0 in step S102, the process proceeds to step S103. In step S103, the inter-view reference vector generation section 83 designates the MV MVbase=l0 of the direction L0 as the predicted vector PMV_L0 of the direction L0 of the target PU. The predicted vector generation process is ended.

[0187]

When it is determined that the MV MVbase = l0 of the direction L0 of the different view (base view) is unavailable in step S101 or when it is determined that POCcurr_l0 is not equal to POCbase_l0 in step S102, the process proceeds to step S104.

[0188]

In step S104, the inter-view reference vector generation section 83 finds the MV memory 36-N in which MVs of different views are stored and determines whether the MV MVbase = l1 of the direction L1 of the different view (base view) is available.

[0189]

When it is determined that the MV MVbase = l1 of the direction L1 of the different view (base view) is available in step S104, the process proceeds to step S105. In step S105, the inter-view reference vector generation section 83 determines whether POCcurr_l0 which is a Ref POC of the target PU is equal to

[0191]

When it is determined that the MV MVbase = l1 of the direction L1 of the

different view (base view) is unavailable in step S104 or when it is determined that POCcurr_l0 is not equal to POCbase_l1 in step S105, the process proceeds to step S107.

[0192]

In step S107, the inter-view reference vector generation section 83 re-determines whether the MV MVbase = l0 of the direction L0 of the different view (base view) is available.

[0193]

When it is determined that the MV MVbase = l0 of the direction L0 of the different view (base view) is available in step S107, the process proceeds to step S108. In step S108, the inter-view reference vector generation section 83 scales the MV MVbase = l0 of the direction L0 according to POCcurr_l0 which is a Ref POC of the target PU and POCbase_l0 which is a Ref POC of the reference PU. Then, the inter-view reference vector generation section 83 designates the scaled MV MVbase = l0 as the predicted vector PMV_L0 of the direction L0 of the target PU, and the predicted vector generation process is ended.

[0194]

When it is determined that the MV MVbase = l0 of the direction L0 of the different view (base view) is unavailable in step S107, the process proceeds to step S109.

[0195]

In step S109, the inter-view reference vector generation section 83 re-

of the target PU and POCbase_l1 which is a Ref POC of the reference PU. Then, the inter-view reference vector generation section 83 designates the scaled MV

MVbase = l1 as the predicted vector PMV_L0 of the direction L0 of the target PU, and the predicted vector generation process is ended.

[0197]

When it is determined that the MV MVbase = l1 of the direction L1 of the different view (base view) is unavailable in step S109, the process proceeds to step S111. In step S111, the inter-view reference vector generation section 83 determines that there is no predicted vector PMV_L0 of the direction L0 of the target PU and the predicted vector generation process is ended.

[0198]

[Example of predicted vector generation process]

Next, the predicted vector generation process of step S84 of FIG. 11 will be described with reference to a flowchart of FIG. 13. Also, in the example of FIG. 13, a predicted vector generation process in the direction L1 is shown.

[0199]

In step S131, the inter-view reference vector generation section 83 finds the MV memory 36-N in which MVs of different views are stored and determines whether the MV MVbase = l1 of the direction L1 of the different view (base view) is available.

[0200]

When it is determined that the MV MVbase = l1 of the direction L1 of the different view (base view) is available in step S131, the process proceeds to step S132. In step S132, the inter-view reference vector generation section 83

vector generation process is ended.

[0202]

When it is determined that the MV MVbase = l1 of the direction L1 of the different view (base view) is unavailable in step S131 or when it is determined that POCcurr_l1 is not equal to POCbase_l1 in step S132, the process proceeds to step S134.

[0203]

In step S134, the inter-view reference vector generation section 83 finds the MV memory 36-N in which MVs of different views are stored and determines whether the MV MVbase = l0 of the direction L0 of the different view (base view) is available.

[0204]

When it is determined that the MV MVbase = l0 of the direction L0 of the different view (base view) is available in step S134, the process proceeds to step S135. In step S135, the inter-view reference vector generation section 83 determines whether POCcurr_l1 which is a Ref POC of the target PU is equal to POCbase_l0 which is a Ref POC of the reference PU.

[0205]

When it is determined that POCcurr_l1 is equal to POCbase_l0 in step S135, the process proceeds to step S136. In step S136, the inter-view reference vector generation section 83 designates the MV MVbase=l0 of the direction L0 as the predicted vector PMV_L1 of the direction L1 of the target PU. Then, the predicted vector generation process is ended.

[0206]

(base view) is available.

[0208]

When it is determined that the MV MVbase = l1 of the direction L1 of the different view (base view) is available in step S137, the process proceeds to step S138. In step S138, the inter-view reference vector generation section 83 scales the MV MVbase = l1 of the direction L1 according to POCcurr_l1 which is a Ref POC of the target PU and POCbase_l1 which is a Ref POC of the reference PU. Then, the inter-view reference vector generation section 83 designates the scaled MVbase = l1 as the predicted vector PMV_L1 of the direction L1 of the target PU, and the predicted vector generation process is ended.

[0209]

When it is determined that the MV MVbase = l1 of the direction L1 of the different view (base view) is unavailable in step S137, the process proceeds to step S139.

[0210]

In step S139, the inter-view reference vector generation section 83 re-determines whether the MV MVbase = l0 of the direction L0 of the different view (base view) is available.

[0211]

When it is determined that the MV MVbase = l0 of the direction L0 of the different view (base view) is available in step S137, the process proceeds to step S140. In step S140, the inter-view reference vector generation section 83 scales the MV MVbase = l0 of the direction L0 according to POCcurr_l1 which is a Ref POC of the target PU and POCbase_l0 which is a Ref POC of the reference PU. Then, determines that there is no predicted vector PMV_L1 of the direction L1 of the target PU and the predicted vector generation process is ended.

[0213]

In this manner, when the Ref POC (Ref 0) of the current PU is different from the Ref POC (Ref 0) of the reference PU in the different view, the MV of the reference PU is scaled and the scaled MV is configured to serve as a candidate for the predicted vector of the current PU.

[0214]

Thereby, it is possible to improve the encoding efficiency of the MV because it is possible to generate a predicted vector having a high correlation.

[0215]

<3. Second embodiment>

[Configuration example of multi-view image decoding device]

FIG. 16 illustrates a configuration of an embodiment of a decoder constituting a multi-view image decoding device serving as an image processing device to which the present disclosure is applied.

[0216]

The multi-view image decoding device, for example, includes decoders 211-1 to 211-M for decoding multi-view images.

[0217]

The decoder 211-1 decodes encoded data corresponding to a color image of a non-base view from an encoded stream encoded by the encoder 11-1 in an HEVC scheme, and generates the color image of the non-base view.

[0218]

In the example of FIG. 16, the decoder 211-1 is configured to include an accumulation buffer 221, a lossless decoding section 222, an inverse quantization

section 223, an inverse orthogonal transform section 224, a calculation section 225, an in-loop filter 226, and a picture reordering buffer 227, and a digital/analog (D/A) transform section 228. In addition, the decoder 211-1 is configured to include a DPB 229-1, an intra-picture prediction section 230, a motion compensation section 231, a predicted image selection section 232, and an MV memory 233-1.

[0220]

The accumulation buffer 221 is a reception section for receiving corresponding encoded data in an encoded stream from the encoder 11-1. The accumulation buffer 221 temporarily stores the received encoded data and supplies the stored data to the lossless decoding section 222. The encoded data includes header information as well as encoded data (quantized residual coefficient) of a color image of a base view.

[0221]

The lossless decoding section 222 restores the quantized residual coefficient or header information by performing variable length decoding on the encoded data from the accumulation buffer 221. Then, the lossless decoding section 222 supplies a quantization value to the inverse quantization section 223 and supplies a corresponding piece of the header information to each of the intra-picture prediction section 230 and the motion compensation section 231.

[0222]

The inverse quantization section 223 inversely quantizes the quantized residual coefficient from the lossless decoding section 222 and supplies the inversely

The calculation section 225 performs decoding by designating a block supplied from the inverse orthogonal transform section 224 as a target block of a

decoding target and adding a predicted image supplied from the predicted image selection section 232 to the target block if necessary. The calculation section 225 supplies a decoded image obtained as a result thereof to the in-loop filter 226.

[0225]

The in-loop filter 226, for example, is constituted of a deblocking filter. Also, for example, when the HEVC scheme is adopted, the in-loop filter 226 is constituted of a deblocking filter and an adaptive offset filter. The in-loop filter 226, for example, performs similar filtering to the in-loop filter 31 of FIG. 4 on the decoded image from the calculation section 225 and supplies the decoded image after the filtering to the picture reordering buffer 227.

[0226]

The picture reordering buffer 227 reorders a picture sequence to an original sequence (display order) by temporarily storing and reading a picture of the decoded image from the in-loop filter 226 and supplies a reordering result to the D/A conversion section 228.

[0227]

When it is necessary to output the picture from the picture reordering buffer 227 in an analog signal, the D/A conversion section 228 performs D/A conversion on the picture and outputs a D/A conversion result.

[0228]

In addition, the in-loop filter 226 supplies the DPB 229-1 with decoded images of an intra (I)-picture, a P-picture, and B-pictures that are referable pictures

predicted image is performed by the calculation section 225) to be performed at a later time is generated. Also, the DPB 229-1 is shared by the decoder 211-M of the

other view.

[0230]

The intra-picture prediction section 230 recognizes whether the target block (PU) is encoded using the predicted image generated in the intra prediction (intra-picture prediction) based on the header information (intra prediction mode) from the lossless decoding section 222.

[0231]

When the target block is encoded using the predicted image generated in intra prediction, the intra-picture prediction section 230 reads an already decoded portion (decoded image) from a picture (target picture) including the target block from the in-loop filter 226 as in the intra-picture prediction section 33 of FIG. 4. Then, the intra-picture prediction section 230 supplies part of the decoded image of the target picture read from the in-loop filter 226 as the predicted image of the target block to the predicted image selection section 232.

[0232]

The motion compensation section 231 recognizes whether the target block is encoded using a predicted image generated in the inter prediction based on header information from the lossless decoding section 222.

[0233]

When the target block is encoded using a predicted image generated in the inter prediction, the motion compensation section 231 recognizes an optimum prediction mode of the target block based on the header information from the lossless

The motion compensation section 231 reads a candidate picture (inter prediction reference picture or inter-view prediction reference picture) corresponding

to a reference image index from candidate pictures stored in the DPB 229-1 or 229-N.

[0236]

Then, the motion compensation section 231 generates a predicted vector for use in decoding of the MV based on an index of the predicted vector from header information from the lossless decoding section 222 in the case of the AMVP mode.

[0237]

For example, when the index of the predicted vector indicates a spatially predicted vector, the motion compensation section 231 generates a predicted vector using spatially neighboring blocks adjacent thereto in the same picture. When the index of the predicted vector indicates a temporally predicted vector, the motion compensation section 231 generates a predicted vector by reading MVs of temporally neighboring blocks corresponding or adjacent thereto in pictures of different times of the same view from the MV memory 233-1. When the index of the predicted vector indicates an inter-view predicted vector, the motion compensation section 231 reads MVs of reference blocks (Cor PU of FIG. 1) in different views of the same time from the MV memory 233-N in which the MVs of the different views are stored and generates a predicted vector.

[0238]

The motion compensation section 231 recognizes the MV representing motion used in generation of the predicted image of the target block by adding motion information of the header information from the lossless decoding section 222 to the generated predicted vector. Then, like the motion prediction/compensation

predicted image.

[0240]

The motion compensation section 231 generates an MV based on a merge index in header information from the lossless decoding section 222 in the case of the M/S mode.

[0241]

For example, when the merge index indicates a spatially predicted vector, the motion compensation section 231 generates the MV using spatially neighboring blocks adjacent thereto in the same picture. When the merge index indicates a temporally predicted vector, the motion compensation section 231 generates an MV by reading MVs of temporally neighboring blocks corresponding or adjacent thereto in pictures of different times of the same view from the MV memory 233-1. When the merge index indicates an inter-view predicted vector, the motion compensation section 231 reads MVs of reference blocks (Cor PU of FIG. 1) in different views of the same time from the MV memory 233-N in which the MVs of the different views are stored and generates an MV.

[0242]

Then, like the motion prediction/compensation section 34 of FIG. 4, the motion compensation section 231 generates a predicted image by performing the motion compensation of the reference picture according to the MV. The motion compensation section 231 supplies the predicted image to the predicted image selection section 232.

[0243]

When the predicted image is supplied from the intra-picture prediction compensation section 231 as (a candidate for) the MV to be referred to when the predicted vector for encoding the MV to be executed at a later time is generated.

Also, the MV memory 233-1 is shared by the decoder 211-M of the other view.

[0245]

Also, the MV memory 233-N is provided in the encoder 11-N having the different view, and stores the MV determined in the decoder 211-N as (the candidate for) the MV to be referred to when the predicted vector for encoding the MV to be executed at a later time is generated. The MV memory 233-N is shared by the motion compensation section 231 or the decoder 211-M of the other view.

[0246]

[Configuration of motion compensation section]

FIG. 15 is a block diagram illustrating a configuration example of the motion compensation section of FIG. 14.

[0247]

In the example of FIG. 15, the motion compensation section 231 is configured to include an automatic reference index generation section 251, an AMVP mode vector prediction section 252, and an M/S mode vector prediction section 253.

[0248]

When the prediction mode is not the inter prediction mode, the merge mode or the skip mode and the merge index of the header information are supplied from the lossless decoding section 222 to the automatic reference index generation section 251.

[0249]

The automatic reference index generation section 251 automatically

lossless decoding section 222 to the AMVP mode vector prediction section 252.

[0251]

The AMVP mode vector prediction section 252 reads a candidate picture (inter prediction reference picture or inter-view prediction reference picture) corresponding to a reference image index from candidate pictures stored in the DPB 229-1 or 229-N according to the inter prediction mode.

[0252]

The AMVP mode vector prediction section 252 generates a predicted vector for use in decoding of the MV based on the index of the predicted vector.

[0253]

For example, when the index of the predicted vector indicates a spatially predicted vector, the AMVP mode vector prediction section 252 generates a predicted vector using spatially neighboring blocks adjacent thereto in the same picture. When the index of the predicted vector indicates a temporally predicted vector, the AMVP mode vector prediction section 252 generates a predicted vector by reading MVs of temporally neighboring blocks corresponding or adjacent thereto in pictures of different times of the same view from the MV memory 233-1. When the index of the predicted vector indicates an inter-view predicted vector, the AMVP mode vector prediction section 252 reads MVs of reference blocks (Cor PU of FIG. 1) in different views of the same time from the MV memory 233-N in which the MVs of the different views are stored, and generates a predicted vector.

[0254]

The AMVP mode vector prediction section 252 recognizes the MV representing motion used in generation of the predicted image of the target block by

prediction reference picture) corresponding to the reference image index from candidate pictures stored in the DPB 229-1 or 229-N.

[0256]

The M/S mode vector prediction section 253 generates the MV based on a merge index in the header information from the lossless decoding section 222.

[0257]

For example, when the merge index indicates a spatially predicted vector, the M/S mode vector prediction section 253 generates an MV using spatially neighboring blocks adjacent thereto in the same picture. When the merge index indicates a temporally predicted vector, the M/S mode vector prediction section 253 generates an MV by reading MVs of corresponding blocks associated by MVs in pictures of different times of the same view from the MV memory 233-1. When the merge index indicates an inter-view predicted vector, the M/S mode vector prediction section 253 reads MVs of reference blocks (Cor PU of FIG. 1) in different views of the same time from the MV memory 233-N in which the MVs of the different views are stored and generates an MV. Information of the generated MV is temporarily saved in the spatial MV memory 262 of FIG. 16 as will be described later.

[0258]

The M/S mode vector prediction section 253 generates a predicted image by performing the motion compensation of the reference picture according to the MV. The generated predicted image is supplied to the predicted image selection section 232.

[0259]

[Configuration of AMVP mode vector prediction section]

[0261]

The predicted image generation section 261 inputs an MV generated by

adding a predicted vector to the MV difference information by the addition section 263 via the spatial MV memory 262. The predicted image generation section 261 reads a reference image corresponding to a reference image index (Ref index) from the lossless decoding section 222 from the DPB 229-1 or 229-N, and generates a predicted image (pred. image) by performing the motion compensation of the read reference image according to the MV. The generated predicted image is supplied to the predicted image selection section 232.

[0262]

The spatial MV memory 262 stores the MV generated by the addition section 263 as the candidate for use in generation of the predicted vector to be performed at a later time. In the spatial MV memory 262, the MV is stored in each unit (PU) of a block in which the MV is obtained. Also, the MV of the M/S mode is also stored in the spatial MV memory 64.

[0263]

The addition section 263 generates the MV by inputting the predicted vector generated by the predicted vector generation section 264 or the predicted vector generation section 265 via the switch 266 and adding the input predicted vector to difference information of the MV supplied from the lossless decoding section 222. The addition section 263 causes the generated MV to be stored in the spatial MV memory 262.

[0264]

The predicted vector generation section 264 generates a spatially predicted

the predicted vector supplied from the lossless decoding section 222 from the MV memory 233-1 or 233-N. The predicted vector generation section 265 supplies the

generated predicted vector to the addition section 263 via the switch 266.

[0266]

That is, when the index of the predicted vector indicates a temporally predicted vector, the predicted vector generation section 265 generates a predicted vector by reading MVs of corresponding blocks associated by MVs in pictures of different times of the same view from the MV memory 233-1. At this time, based on POC information from the POC transform section 267, scaling is performed on the MV of the corresponding block when the reference POC (Ref POC) of the target block is different from the reference POC (Ref POC) of the corresponding block. That is, the scaled MV serves as a predicted vector.

[0267]

When the index of the predicted vector indicates an inter-view predicted vector, the AMVP mode vector prediction section 252 reads MVs of reference blocks (Cor PU of FIG. 1) in different views of the same time from the MV memory 233-N in which the MVs of the different views are stored and generates a predicted vector. At this time, based on POC information from the POC transform section 267, scaling is performed on the MV of the reference block when the reference POC (Ref POC) of the target block is different from a reference POC (Ref POC) of the reference block. That is, the scaled MV serves as the predicted vector.

[0268]

The POC transform section 267 transforms the reference image index (Ref index) of the target block from the lossless decoding section 222 into a POC and

[0270]

In the example of FIG. 16, the predicted vector generation section 265 is

configured to include an intra-view reference vector generation section 281 and an inter-view reference vector generation section 282.

[0271]

When the index of the predicted vector indicates a temporally (TMVP) predicted vector, the index (MV index) of the predicted vector is supplied from the lossless decoding section 222 to the intra-view reference vector generation section 281.

[0272]

The intra-view reference vector generation section 281 generates a predicted vector by reading an MV of a corresponding block indicated by an index of the predicted vector, that is, associated by the MV, in pictures of different times of the same view from the MV memory 233-1.

[0273]

Also, at this time, based on POC information from the POC transform section 267, scaling is performed on an MV of a corresponding block when the reference POC (Ref POC) of the target block is different from a reference POC (Ref POC) of the corresponding block. That is, the scaled MV serves as the predicted vector.

[0274]

The intra-view reference vector generation section 281 supplies the generated predicted vector to the addition section 263 via the switch 266.

[0275]

disparity vector from MVs of neighboring blocks adjacent to the target block from the spatial MV memory 262 and obtains reference blocks in different views of the

same time based on the found disparity vector. Then, the inter-view reference vector generation section 282 generates a predicted vector by reading an MV of the reference block indicated by the index of the predicted vector from the MV memory 233-N in which MVs of different views are stored.

[0277]

Also, at this time, based on POC information from the POC transform section 267, scaling is performed on an MV of a reference block when the reference POC (Ref POC) of the target block is different from a reference POC (Ref POC) of the reference block. That is, the scaled MV serves as the predicted vector.

[0278]

The inter-view reference vector generation section 282 supplies the generated predicted vector to the addition section 263 via the switch 266.

[0279]

[Operation of decoder]

Next, a decoding process of the decoder 211-1 of FIG. 14 will be described with reference to the flowchart of FIG. 18. Also, the decoders 211-N and 211-M for decoding images of other views perform similar decoding processes.

[0280]

The accumulation buffer 221 temporarily stores encoded data corresponding to a received color image of a non-base view and supplies the stored encoded data to the lossless decoding section 222.

[0281]

[0283]

In step S213, the inverse orthogonal transform section 224 performs an

inverse orthogonal transform on the transform coefficient from the inverse quantization section 223 and supplies an inverse orthogonal transform result to the calculation section 225.

[0284]

In step S214, the intra-picture prediction section 230 determines whether prediction for a target block (PU) is intra-picture prediction based on header information (intra prediction mode) from the lossless decoding section 222. When it is determined that the prediction is the intra-picture prediction in step S214, the process proceeds to step S215. In step S215, the intra-picture prediction section 230 performs the intra-picture prediction.

[0285]

When it is determined that the prediction is not the intra-picture prediction in step S214, the process proceeds to step S216. In step S216, the motion compensation section 231 performs a motion compensation process. This motion compensation process will be described later with reference to FIG. 19.

[0286]

In the process of step S216, a predicted vector according to the predicted vector index is generated when the motion prediction is a motion prediction mode and an MV is generated. In addition, a reference image according to the reference image index is read, motion compensation is performed according to the generated MV, and a predicted image is generated.

[0287]

[0289]

When the predicted image is supplied from the intra-picture prediction

section 230, the predicted image selection section 232 selects the predicted image and supplies the selected predicted image to the calculation section 225. When the predicted image is supplied from the motion compensation section 231, the predicted image selection section 232 selects the predicted image and supplies the selected predicted image to the calculation section 225.

[0290]

In step S218, the calculation section 225 adds a block (difference) supplied from the inverse orthogonal transform section 224 to the predicted image supplied from the predicted image selection section 232. The calculation section 225 supplies a decoded image obtained as a result thereof to the in-loop filter 226.

[0291]

In step S219, the in-loop filter 226 determines whether an LCU ends. When it is determined that the LCU does not end in step S219, the process returns to step S211 and the process subsequent thereto is iterated.

[0292]

When it is determined that the LCU ends in step S219, the process proceeds to step S220. In step S220, the in-loop filter 226 removes (reduces) block distortion occurring in the decoded image by filtering the decoded image from the calculation section 225.

[0293]

In step S221, the in-loop filter 226 saves the decoded image after the filtering in the DPB 229-1.

Then, the motion compensation section 231 saves the compressed MV in the MV memory 233-1.

[0296]

In step S224, the decoder 211-M of the other view decodes a picture of the other view. Also, the decoding process is basically similar to the decoding process of FIG. 18.

[0297]

As described above, the decoding process is performed.

[0298]

[Example of motion compensation process]

Next, the motion compensation process of step S216 of FIG. 18 will be described with reference to a flowchart of FIG. 19.

[0299]

The lossless decoding section 222 decodes the motion prediction mode in the header information in step S241, and determines whether the prediction mode is an inter prediction mode in step S242.

[0300]

When it is determined that the prediction mode is the inter prediction mode in step S242, the lossless decoding section 222 supplies the inter prediction mode (inter mode), a reference image index (Ref index), MV difference information (Mvd), and an index (MV index) of the predicted vector to the AMVP mode vector prediction section 252. Then, the process proceeds to step S243.

[0301]

In step S243, the AMVP mode vector prediction section 252 performs

supplied to the predicted image selection section 232.

[0303]

On the other hand, when it is determined that the mode is not the inter prediction mode in step S242, the lossless decoding section 222 supplies the merge mode or the skip mode and a merge index to the automatic reference index generation section 251. Then, the process proceeds to step S244.

[0304]

In step S244, the automatic reference index generation section 251 automatically generates a reference image index and supplies the generated reference image index (Ref index) and merge index to the M/S mode vector prediction section 253 along with the merge mode or the skip mode from the lossless decoding section 222.

[0305]

In step S245, the M/S mode vector prediction section 253 performs a vector prediction process of the merge mode or the skip mode. That is, the M/S mode vector prediction section 253 reads a candidate picture (inter prediction reference picture) corresponding to the reference image index from candidate pictures stored in the DPB 229-1 or 229-N.

[0306]

The M/S mode vector prediction section 253 generates an MV based on a merge index in the header information from the lossless decoding section 222.

[0307]

For example, when the merge index indicates a spatially predicted vector, the M/S mode vector prediction section 253 generates an MV using spatially

are stored and generates an MV.

[0308]

The M/S mode vector prediction section 253 generates a predicted image by performing motion compensation of a reference picture according to the MV. The generated predicted image is supplied to the predicted image selection section 232.

[0309]

[Vector prediction process of AMVP mode]

Next, the vector prediction process of the AMVP mode will be described with reference to a flowchart of FIG. 20.

[0310]

In step S261, the lossless decoding section 222 decodes MV difference information (MVd) of the header information and supplies the decoded MV difference information to the addition section 263.

[0311]

In step S262, the lossless decoding section 222 decodes a reference image index of the header information and supplies the decoded reference image index (Ref index) to the predicted image generation section 261 and the POC transform section 267.

[0312]

In step S263, the lossless decoding section 222 decodes an index of a predicted vector of the header information.

[0313]

In step S264, the lossless decoding section 222 refers to the predicted vector index decoded in step S263 and determines whether the predicted vector is spatial.

generates the spatially predicted vector by reading an MV indicated by the index of the predicted vector supplied from the lossless decoding section 222 from the spatial

MV memory 262. The predicted vector generation section 264 supplies the generated predicted vector to the addition section 263 via the switch 266.

[0316]

When it is determined that the predicted vector is not spatial in step S264, the process proceeds to step S266.

[0317]

In step S266, the predicted vector generation section 265 generates a non-spatially predicted vector. A process of generating the non-spatially predicted vector will be described later with reference to FIG. 21.

[0318]

In the process of step S266, the index of the predicted vector is supplied from the lossless decoding section 222, an MV indicated by the index of the predicted vector is read from the MV memory 233-1 or 233-N, and a non-spatially (that is, TMVP or IVMP) predicted vector is generated. The generated predicted vector is supplied to the addition section 263 via the switch 266.

[0319]

In step S267, the addition section 263 generates the MV. That is, the predicted vector generated by the predicted vector generation section 264 or 265 is input to the addition section 263 via the switch 266. The addition section 263 generates the MV by adding the input predicted vector to difference information of the MV supplied from the lossless decoding section 222.

[0320]

the lossless decoding section 222 from the DPB 229-1 or 229-N. According to the MV from the spatial MV memory 262, the predicted image generation section 261

generates a predicted image by performing motion compensation of a read reference image according to the MV from the spatial MV memory 262.

[0322]

[Process of generating non-spatially predicted vector]

Next, the process of generating the non-spatially predicted vector in step S266 of FIG. 20 will be described with reference to the flowchart of FIG. 21.

[0323]

In step S281, the lossless decoding section 222 refers to the index of the predicted vector decoded in step S263 of FIG. 20 and determines whether the predicted vector is temporal. When it is determined that the predicted vector is temporal in step S281, the lossless decoding section 222 supplies the index of the predicted vector to the intra-view reference vector generation section 281. Then, the process proceeds to step S282.

[0324]

In step S282, the intra-view reference vector generation section 281 generates a predicted vector by the TMVP. That is, the intra-view reference vector generation section 281 generates the predicted vector by reading an MV of a corresponding block indicated by the index of the predicted vector, that is, associated by the MV, in pictures of different times of the same view from the MV memory 233-1. The generated predicted vector is supplied to the addition section 263 via the switch 266.

[0325]

[0327]

That is, in step S283, the inter-view reference vector generation section 282

finds a disparity vector from MVs of neighboring blocks adjacent to the target block (PU) from the spatial MV memory 262 and calculates a disparity based on the found disparity vector.

[0328]

In step S284, the inter-view reference vector generation section 282 selects a PU of a position shifted by the disparity obtained in step S283 as the reference PU in a different view.

[0329]

In step S285, the inter-view reference vector generation section 282 generates a predicted vector from an MV of a selected reference PU by reading the MV of the selected reference PU from the MV memory 233-N in which MVs of the different views are stored. Because the predicted vector generation process is basically the same process as the predicted vector generation process described above with reference to FIGS. 12 and 13, redundant description thereof is omitted.

[0330]

That is, it is determined whether the reference POC (Ref POC) of the target block is different from a reference POC (Ref POC) of a reference block based on POC information from the POC transform section 267 in step S285. When it is determined that the Ref POCs are different, scaling is performed on the MV of the reference block. That is, when it is determined that the Ref POCs are different, an MV of the reference block is scaled and a predicted vector is generated.

[0331]

of the reference block. That is, the MV of the reference block of the different view can also be designated as the candidate of the predicted vector. Therefore, there is a

significant effect in the improvement of encoding efficiency because MVs having a high correlation can be scaled and used.

[0333]

Also, although the case of the AMVP mode has been described above in detail, the present technology is also applicable to the merge mode. Also, in the case of the merge mode, as in the case of the TMVP, the Ref index is fixed to 0 and the MV of the reference PU is scaled when the Ref POC of the reference PU of the base view is different from the Ref POC of the current PU and the scaled MV serves as a predicted vector.

[0334]

In this case, a processing circuit of the TMVP and the IVMP can be made common.

[0335]

In addition, an example in which an inter-MV of the reference block shifted by a disparity indicated by a disparity vector of a block adjacent to the target block is used after being scaled in the time direction according to the POC in a different view from the target block when the predicted vector of the inter-MV of the target block is obtained has been described above.

[0336]

On the other hand, the present technology is also applicable when the inter-view MV is used as the predicted vector. That is, when an MV of a corresponding block of a different time corresponding to a target block of a certain time is an inter-

[0338]

Also, the present disclosure, for example, is applicable to an image

encoding device and an image decoding device to be used when image information (bitstream) compressed by an orthogonal transform such as a discrete cosine transform and motion compensation as in the HEVC scheme or the like is received via network media such as satellite broadcasting, a cable television, the Internet, and a mobile phone. In addition, the present disclosure is applicable to an image encoding device and an image decoding device to be used when processing is performed on storage media such as an optical disc, a magnetic disc and a flash memory.

[0339]

Also, the present technology is applicable, for example, to HTTP streaming such as MPEG DASH for selecting and using appropriate encoded data from a plurality of pieces of encoded data in which prepared resolutions and the like are different from each other in units of segments.

[0340]

<4. Third embodiment>

[Configuration example of computer]

The above described series of processes can be executed by hardware or can be executed by software. When the series of processes are to be performed by software, the programs forming the software are installed into a computer. Here, a computer includes a computer which is incorporated in dedicated hardware or a general-purpose personal computer (PC) which can execute various functions by installing various programs into the computer, for example.

by a bus 804.

[0343]

An input and output interface (I/F) 805 is further connected to the bus 804. An input section 806, an output section 807, a storage section 808, a communication section 809, and a drive 810 are connected to the input and output I/F 805.

[0344]

The input section 806 is formed with a keyboard, a mouse, a microphone, and the like. The output section 807 is formed with a display, a speaker, and the like. The storage section 808 is formed with a hard disk, a nonvolatile memory, or the like. The communication section 809 is formed with a network interface or the like. The drive 810 drives a removable medium 811 such as a magnetic disk, an optical disk, a magneto-optical disk, or a semiconductor memory.

[0345]

In the computer configured as described above, the CPU 801 loads the programs stored in the storage section 808 into the RAM 803 via the input and output I/F 805 and the bus 804, and executes the programs, so that the above described series of processes are performed.

[0346]

The program executed by the computer 800 (the CPU 801) may be provided by being recorded on the removable medium 811 as a packaged medium or the like. The program can also be provided via a wired or wireless transfer medium, such as a local area network, the Internet, or a digital satellite broadcast.

[0347]

In the computer, by loading the removable medium 811 into the drive 810,

program that is processed in time series according to the sequence described in this specification or a program that is processed in parallel or at necessary timing such as

upon calling.

[0349]

In the present disclosure, steps of describing the program to be recorded on the recording medium may include processing performed in time-series according to the description order and processing not processed in time-series but performed in parallel or individually.

[0350]

In the specification, the system refers to the entire apparatus including a plurality of devices (apparatuses).

[0351]

Further, an element described as a single device (or processing unit) above may be divided and configured as a plurality of devices (or processing units). On the contrary, elements described as a plurality of devices (or processing units) above may be configured collectively as a single device (or processing unit). Further, an element other than those described above may be added to each device (or processing unit). Furthermore, a part of an element of a given device (or processing unit) may be included in an element of another device (or another processing unit) as long as the configuration or operation of the system as a whole is substantially the same. In other words, an embodiment of the disclosure is not limited to the embodiments described above, and various changes and modifications may be made without departing from the scope of the disclosure.

[0352]

[0353]

<5. Applications>

[First Application: Television Receivers]

FIG. 23 illustrates an example of a schematic configuration of a television device to which the embodiment is applied. A television device 900 includes an antenna 901, a tuner 902, a demultiplexer 903, a decoder 904, an video signal processing section 905, a display section 906, an audio signal processing section 907, a speaker 908, an external I/F 909, a control section 910, a user I/F 911, and a bus 912.

[0354]

The tuner 902 extracts a signal of a desired channel from broadcast signals received via the antenna 901, and demodulates the extracted signal. The tuner 902 then outputs an encoded bit stream obtained through the demodulation to the demultiplexer 903. That is, the tuner 902 serves as a transmission means of the television device 900 for receiving an encoded stream in which an image is encoded.

[0355]

The demultiplexer 903 demultiplexes the encoded bit stream to obtain a video stream and an audio stream of a program to be viewed, and outputs each stream obtained through the demultiplexing to the decoder 904. The demultiplexer 903 also extracts auxiliary data such as electronic program guides (EPGs) from the encoded bit stream, and supplies the extracted data to the control section 910. Additionally, the demultiplexer 903 may perform descrambling when the encoded bit stream has been scrambled.

[0356]

from the decoder 904, and causes the display section 906 to display the video. The video signal processing section 905 may also cause the display section 906 to display

an application screen supplied via a network. Further, the video signal processing section 905 may perform an additional process such as noise removal (suppression), for example, on the video data in accordance with the setting. Furthermore, the video signal processing section 905 may generate an image of a graphical user I/F (GUI) such as a menu, a button and a cursor, and superimpose the generated image on an output image.

[0358]

The display section 906 is driven by a drive signal supplied from the video signal processing section 905, and displays a video or an image on a video screen of a display device (e.g. liquid crystal display, plasma display, organic electrioluminescence display(OLED), etc.).

[0359]

The audio signal processing section 907 performs a reproduction process such as D/A conversion and amplification on the audio data input from the decoder 904, and outputs a sound from the speaker 908. The audio signal processing section 907 may also perform an additional process such as noise removal (suppression) on the audio data.

[0360]

The external I/F 909 is an I/F for connecting the television device 900 to an external device or a network. For example, a video stream or an audio stream received via the external I/F 909 may be decoded by the decoder 904. That is, the external I/F 909 also serves as a transmission means of the television device 900 for the television device 900, for example. The CPU controls the operation of the television device 900, for example, in accordance with an operation signal input from

the user I/F 911 by executing the program.

[0362]

The user I/F 911 is connected to the control section 910. The user I/F 911 includes, for example, a button and a switch used for a user to operate the television device 900, and a receiving section for a remote control signal. The user I/F 911 detects an operation of a user via these structural elements, generates an operation signal, and outputs the generated operation signal to the control section 910.

[0363]

The bus 912 connects the tuner 902, the demultiplexer 903, the decoder 904, the video signal processing section 905, the audio signal processing section 907, the external I/F 909, and the control section 910 to each other.

[0364]

The decoder 904 has a function of the image decoding device 60 according to the embodiment in the television device 900 configured in this manner. It is possible to improve the encoding efficiency of encoding or decoding of an MV in a multi-view image.

[0365]

[Second Application: Mobile phones]

FIG. 24 illustrates an example of a schematic configuration of a mobile phone to which the embodiment is applied. A mobile phone 920 includes an antenna 921, a communication section 922, an audio codec 923, a speaker 924, a microphone 925, a camera section 926, an image processing section 927, a

image processing section 927, the demultiplexing section 928, the recording/reproduction section 929, the display section 930, and the control section

931 to each other.

[0367]

The mobile phone 920 performs an operation such as transmission and reception of an audio signal, transmission and reception of email or image data, image capturing, and recording of data in various operation modes including an audio call mode, a data communication mode, an image capturing mode, and a videophone mode.

[0368]

An analogue audio signal generated by the microphone 925 is supplied to the audio codec 923 in the audio call mode. The audio codec 923 converts the analogue audio signal into audio data, has the converted audio data subjected to the A/D conversion, and compresses the converted data. The audio codec 923 then outputs the compressed audio data to the communication section 922. The communication section 922 encodes and modulates the audio data, and generates a transmission signal. The communication section 922 then transmits the generated transmission signal to a base station (not illustrated) via the antenna 921. The communication section 922 also amplifies a wireless signal received via the antenna 921 and converts the frequency of the wireless signal to acquire a received signal. The communication section 922 then demodulates and decodes the received signal, generates audio data, and outputs the generated audio data to the audio codec 923. The audio codec 923 extends the audio data, has the audio data subjected to the D/A conversion, and generates an analogue audio signal. The audio codec 923 then

and outputs the generated email data to the communication section 922. The communication section 922 encodes and modulates the email data, and generates a

transmission signal. The communication section 922 then transmits the generated transmission signal to a base station (not illustrated) via the antenna 921. The communication section 922 also amplifies a wireless signal received via the antenna 921 and converts the frequency of the wireless signal to acquire a received signal. The communication section 922 then demodulates and decodes the received signal to restore the email data, and outputs the restored email data to the control section 931. The control section 931 causes the display section 930 to display the content of the email, and also causes the storage medium of the recording/reproduction section 929 to store the email data.

[0370]

The recording/reproduction section 929 includes a readable and writable storage medium. For example, the storage medium may be a built-in storage medium such as RAM and flash memory, or an externally mounted storage medium such as hard disks, magnetic disks, magneto-optical disks, optical discs, unallocated space bitmap (USB) memory, and memory cards.

[0371]

Furthermore, the camera section 926, for example, captures an image of a subject to generate image data, and outputs the generated image data to the image processing section 927 in the image capturing mode. The image processing section 927 encodes the image data input from the camera section 926, and causes the storage medium of the storage/reproduction section 929 to store the encoded stream.

[0372]

wireless signal received via the antenna 921 and converts the frequency of the wireless signal to acquire a received signal. These transmission signal and received

signal may include an encoded bit stream. The communication section 922 then demodulates and decodes the received signal to restore the stream, and outputs the restored stream to the demultiplexing section 928. The demultiplexing section 928 demultiplexes the input stream to obtain a video stream and an audio stream, and outputs the video stream to the image processing section 927 and the audio stream to the audio codec 923. The image processing section 927 decodes the video stream, and generates video data. The video data is supplied to the display section 930, and a series of images is displayed by the display section 930. The audio codec 923 extends the audio stream, has the audio stream subjected to the D/A conversion, and generates an analogue audio signal. The audio codec 923 then supplies the generated audio signal to the speaker 924, and causes a sound to be output.

[0373]

The image processing section 927 has a function of the image encoding device and the image decoding device according to the embodiment in the mobile phone 920 configured in this manner. It is possible to improve the encoding efficiency of encoding or decoding of an MV in a multi-view image.

[0374]

[Third Application: Recording/reproduction device]

FIG. 25 illustrates an example of a schematic configuration of a recording/reproduction device to which the embodiment is applied. A recording/reproduction device 940, for example, encodes audio data and video data of a received broadcast program and records the encoded audio data and the encoded

[0375]

The recording/reproduction device 940 includes a tuner 941, an external I/F

942, an encoder 943, a hard disk drive (HDD) 944, a disc drive 945, a selector 946, a decoder 947, an on-screen display (OSD) 948, a control section 949, and a user I/F 950.

[0376]

The tuner 941 extracts a signal of a desired channel from broadcast signals received via an antenna (not shown), and demodulates the extracted signal. The tuner 941 then outputs an encoded bit stream obtained through the demodulation to the selector 946. That is, the tuner 941 serves as a transmission means of the recording/reproduction device 940.

[0377]

The external I/F 942 is an I/F for connecting the recording/reproduction device 940 to an external device or a network. For example, the external I/F 942 may be an Institute of Electrical and Electronics Engineers (IEEE) 1394 I/F, a network I/F, an USB I/F, a flash memory I/F, or the like. For example, video data and audio data received via the external I/F 942 are input to the encoder 943. That is, the external I/F 942 serves as a transmission means of the recording/reproduction device 940.

[0378]

When the video data and the audio data input from the external I/F 942 have not been encoded, the encoder 943 encodes the video data and the audio data. The encoder 943 then outputs an encoded bit stream to the selector 946.

[0379]

example, a DVD disc (DVD-Video, DVD-RAM, DVD-R, DVD-RW, a DVD+R, DVD+RW, etc.), a Blu-ray (registered trademark) disc, or the like.

[0381]

The selector 946 selects, at the time of recording a video or a sound, an encoded bit stream input from the tuner 941 or the encoder 943, and outputs the selected encoded bit stream to the HDD 944 or the disc drive 945. The selector 946 also outputs, at the time of reproducing a video or a sound, an encoded bit stream input from the HDD 944 or the disc drive 945 to the decoder 947.

[0382]

The decoder 947 decodes the encoded bit stream, and generates video data and audio data. The decoder 947 then outputs the generated video data to the OSD 948. The decoder 904 also outputs the generated audio data to an external speaker.

[0383]

The OSD 948 reproduces the video data input from the decoder 947, and displays a video. The OSD 948 may also superimpose an image of a GUI such as a menu, a button, and a cursor on a displayed video.

[0384]

The control section 949 includes a processor such as a CPU, and a memory such as RAM and ROM. The memory stores a program to be executed by the CPU, program data, and the like. For example, a program stored in the memory is read out and executed by the CPU at the time of activation of the recording/reproduction device 940. The CPU controls the operation of the recording/reproduction device 940, for example, in accordance with an operation signal input from the user I/F 950 by executing the program.

[0386]

The encoder 943 has a function of the image encoding device according to

the embodiment in the recording/reproduction device 940 configured in this manner. The decoder 947 also has a function of the image decoding device according to the embodiment. It is possible to improve the encoding efficiency of encoding or decoding of an MV in a multi-view image.

[0387]

[Fourth Application: Image capturing device]

FIG. 26 illustrates an example of a schematic configuration of an image capturing device to which the embodiment is applied. An image capturing device 960 captures an image of a subject to generate an image, encodes the image data, and records the image data in a recording medium.

[0388]

The image capturing device 960 includes an optical block 961, an image capturing section 962, a signal processing section 963, an image processing section 964, a display section 965, an external I/F 966, a memory 967, a media drive 968, an OSD 969, a control section 970, a user I/F 971, and a bus 972.

[0389]

The optical block 961 is connected to the image capturing section 962. The image capturing section 962 is connected to the signal processing section 963. The display section 965 is connected to the image processing section 964. The user I/F 971 is connected to the control section 970. The bus 972 connects the image processing section 964, the external I/F 966, the memory 967, the media drive 968, the OSD 969, and the control section 970 to each other.

signal through photoelectric conversion. The image capturing section 962 then outputs the image signal to the signal processing section 963.

[0391]

The signal processing section 963 performs various camera signal processes such as knee correction, gamma correction, and color correction on the image signal input from the image capturing section 962. The signal processing section 963 outputs the image data subjected to the camera signal process to the image processing section 964.

[0392]

The image processing section 964 encodes the image data input from the signal processing section 963, and generates encoded data. The image processing section 964 then outputs the generated encoded data to the external I/F 966 or the media drive 968. The image processing section 964 also decodes encoded data input from the external I/F 966 or the media drive 968, and generates image data. The image processing section 964 then outputs the generated image data to the display section 965. The image processing section 964 may also output the image data input from the signal processing section 963 to the display section 965, and cause the image to be displayed. Furthermore, the image processing section 964 may superimpose data for display acquired from the OSD 969 on an image to be output to the display section 965.

[0393]

The OSD 969 generates an image of a GUI such as a menu, a button, and a cursor, and outputs the generated image to the image processing section 964.

[0394]

such as a LAN and the Internet. That is, the external I/F 966 serves as a transmission means of the image capturing device 960.

[0395]

A recording medium to be mounted on the media drive 968 may be a readable and writable removable medium such as magnetic disks, magneto-optical disks, optical discs, and semiconductor memory. The recording medium may also be fixedly mounted on the media drive 968, configuring a non-transportable storage section such as built-in hard disk drives or a solid state drives (SSDs).

[0396]

The control section 970 includes a processor such as a CPU, and a memory such as RAM and ROM. The memory stores a program to be executed by the CPU, program data, and the like. A program stored in the memory is read out and executed by the CPU, for example, at the time of activation of the image capturing device 960. The CPU controls the operation of the image capturing device 960, for example, in accordance with an operation signal input from the user I/F 971 by executing the program.

[0397]

The user I/F 971 is connected to the control section 970. The user I/F 971 includes, for example, a button, a switch, and the like used for a user to operate the image capturing device 960. The user I/F 971 detects an operation made by a user via these structural elements, generates an operation signal, and outputs the generated operation signal to the control section 970.

[0398]

The image processing section 964 has a function of the image encoding video coding (hierarchical coding) is performed, will be described. The scalable video coding, for example, is used for selection of data to be transmitted as examples

illustrated in FIG. 27.

[0400]

In a data transmission system 1000 illustrated in FIG. 27, a distribution server 1002 reads scalable coded data stored in a scalable coded data storage section 1001, and distributes the scalable coded data to a terminal device such as a PC 1004, an AV device 1005, a tablet device 1006, or a mobile phone 1007 via a network 1003.

[0401]

At this time, the distribution server 1002 selects and transmits coded data having proper quality according to capability of the terminal device, communication environment, or the like. Even when the distribution server 1002 transmits unnecessarily high-quality data, a high-quality image is not necessarily obtainable in the terminal device and it may be a cause of occurrence of a delay or an overflow. In addition, a communication band may be unnecessarily occupied or a load of the terminal device may be unnecessarily increased. In contrast, even when the distribution server 1002 transmits unnecessarily low quality data, an image with a sufficient quality may not be obtained. Thus, the distribution server 1002 appropriately reads and transmits the scalable coded data stored in the scalable coded data storage section 1001 as the coded data having a proper quality according to the capability of the terminal device, the communication environment, or the like.

[0402]

For example, the scalable coded data storage section 1001 is configured to store scalable coded data (BL+EL) 1011 in which the scalable video coding is

respect to the PC 1004 or the tablet device 1006 having high processing capability, the distribution server 1002 reads the scalable coded data (BL+EL) 1011 from the

scalable coded data storage section 1001, and transmits the scalable coded data (BL+EL) 1011 without change. On the other hand, for example, with respect to the AV device 1005 or the mobile phone 1007 having low processing capability, the distribution server 1002 extracts the data of the base layer from the scalable coded data (BL+EL) 1011, and transmits the extracted data of the base layer as low quality scalable coded data (BL) 1012 that is data having the same content as the scalable coded data (BL+EL) 1011 but has lower quality than the scalable coded data (BL+EL) 1011.

[0404]

Because an amount of data can easily be adjusted by employing the scalable coded data, the occurrence of the delay or the overflow can be suppressed or the unnecessary increase of the load of the terminal device or the communication media can be suppressed. In addition, because a redundancy between the layers is reduced in the scalable coded data (BL+EL) 1011, it is possible to further reduce the amount of data than when the coded data of each layer is treated as the individual data. Therefore, it is possible to more efficiently use the storage region of the scalable coded data storage section 1001.

[0405]

Because various devices such as the PC 1004 to the mobile phone 1007 are applicable as the terminal device, the hardware performance of the terminal devices differs according to the device. In addition, because there are various applications which are executed by the terminal device, the software performance thereof also

the terminal device which is the data transmission destination before starting the data transmission, and then obtain information related to the terminal device performance

such as hardware performance of the terminal device, or the application (software) performance which is executed by the terminal device, and information related to the communication environment such as an available bandwidth of the network 1003. Then, distribution server 1002 may select an appropriate layer based on the obtained information.

[0407]

Also, the extraction of the layer may be performed in the terminal device. For example, the PC 1004 may decode the transmitted scalable coded data (BL+EL) 1011 and display the image of the base layer or display the image of the enhancement layer. In addition, for example, the PC 1004 may be configured to extract the scalable coded data (BL) 1012 of the base layer from the transmitted scalable coded data (BL+EL) 1011, store the extracted scalable coded data (BL) 1012 of the base layer, transmit to another device, or decode and display the image of the base layer.

[0408]

Of course, the number of the scalable coded data storage sections 1001, the distribution servers 1002, the networks 1003, and the terminal devices are optional. In addition, although the example of the distribution server 1002 transmitting the data to the terminal device is described above, the example of use is not limited thereto. The data transmission system 1000 is applicable to any system which selects and transmits an appropriate layer according to the capability of the terminal device, the communication environment, or the like when the scalable coded data is

In addition, the scalable video coding, for example, is used for transmission via a plurality of communication media as in an example illustrated in FIG. 28.

[0411]

In a data transmission system 1100 illustrated in FIG. 28, a broadcasting station 1101 transmits scalable coded data (BL) 1121 of the base layer by terrestrial broadcasting 1111. In addition, the broadcasting station 1101 transmits scalable coded data (EL) 1122 of the enhancement layer via any arbitrary network 1112 made of a communication network that is wired, wireless, or both (for example, the data is packetized and transmitted).

[0412]

A terminal device 1102 has a function of receiving the terrestrial broadcasting 1111 that is broadcast by the broadcasting station 1101 and receives the scalable coded data (BL) 1121 of the base layer transmitted via the terrestrial broadcasting 1111. In addition, the terminal device 1102 further has a communication function by which the communication is performed via the network 1112, and receives the scalable coded data (EL) 1122 of the enhancement layer transmitted via the network 1112.

[0413]

For example, according to a user's instruction or the like, the terminal device 1102 decodes the scalable coded data (BL) 1121 of the base layer acquired via the terrestrial broadcasting 1111, thereby obtaining or storing the image of the base layer or transmitting the image of the base layer to other devices.

[0414]

In addition, for example, according to the user's instruction, the terminal

As described above, the scalable coded data, for example, can be transmitted via the different communication medium for each layer. Therefore, it is

possible to disperse the load and suppress the occurrence of the delay or the overflow.

[0416]

In addition, according to the situation, the communication medium used for the transmission for each layer may be configured to be selected. For example, the scalable coded data (BL) 1121 of the base layer in which the amount of data is comparatively large may be transmitted via the communication medium having a wide bandwidth, and the scalable coded data (EL) 1122 of the enhancement layer in which the amount of data is comparatively small may be transmitted via the communication media having a narrow bandwidth. In addition, for example, whether the communication medium that transmits the scalable coded data (EL) 1122 of the enhancement layer is the network 1112 or the terrestrial broadcasting 1111 may be switched according to the available bandwidth of the network 1112. Of course, the same is true for data of an arbitrary layer.

[0417]

By controlling in this way, it is possible to further suppress the increase of the load in the data transmission.

[0418]

Of course, the number of the layers is optional, and the number of communication media used in the transmission is also optional. In addition, the number of terminal devices 1102 which are the destination of the data distribution is also optional. Further, although the example of the broadcasting from the broadcasting station 1101 has been described above, the use example is not limited

[0420]

[Third system]

In addition, the scalable video coding is used in the storage of the coded data as an example illustrated in FIG. 29.

[0421]

In an image capturing system 1200 illustrated in FIG. 29, an image capturing device 1201 performs scalable video coding on image data obtained by capturing an image of a subject 1211, and supplies a scalable video result as the scalable coded data (BL+EL) 1221 to a scalable coded data storage device 1202.

[0422]

The scalable coded data storage device 1202 stores the scalable coded data (BL+EL) 1221 supplied from the image capturing device 1201 in quality according to the situation. For example, in the case of normal circumstances, the scalable coded data storage device 1202 extracts data of the base layer from the scalable coded data (BL+EL) 1221, and stores the extracted data as scalable coded data (BL) 1222 of the base layer having a small amount of data at low quality. On the other hand, for example, in the case of notable circumstances, the scalable coded data storage device 1202 stores the scalable coded data (BL+EL) 1221 having a large amount of data at high quality without change.

[0423]

In this way, because the scalable coded data storage device 1202 can save the image at high quality only in a necessary case, it is possible to suppress the decrease of the value of the image due to the deterioration of the image quality and suppress the increase of the amount of data, and it is possible to improve the use

other hand, because the content of the imaged image is likely to be important when a monitoring target is shown as the subject 1211 in the imaged image (in the case of

the notable circumstances), the priority is on the image quality, and the image data (scalable coded data) is stored at high quality.

[0425]

For example, whether the case is the case of the normal circumstances or the notable circumstances may be determined by the scalable coded data storage device 1202 by analyzing the image. In addition, the image capturing device 1201 may be configured to make the determination and transmit the determination result to the scalable coded data storage device 1202.

[0426]

A determination criterion of whether the case is the case of the normal circumstances or the notable circumstances is optional and the content of the image which is the determination criterion is optional. Of course, a condition other than the content of the image can be designated as the determination criterion. For example, switching may be configured to be performed according to the magnitude or waveform of recorded sound, by a predetermined time interval, or by an external instruction such as the user's instruction.

[0427]

In addition, although the two states of the normal circumstances and the notable circumstances have been described above, the number of the states is optional, and for example, switching may be configured to be performed among three or more states such as normal circumstances, slightly notable circumstances, notable circumstances, and highly notable circumstances. However, the upper limit

quality and supply the data to the scalable coded data storage device 1202. In addition, for example, in the case of the notable circumstances, the image capturing

device 1201 may generate the scalable coded data (BL+EL) 1221 of the base layer having a large amount of data at high quality and supply the data to the scalable coded data storage device 1202.

[0429]

Although the monitoring camera has been described above as the example, the usage of the image capturing system 1200 is optional and is not limited to the monitoring camera.

[0430]

Even in the image capturing system 1200 as in FIG. 29, it is possible to obtain effects similar to those described above with reference to FIGS. 1 to 21 by applying the present technology described above with reference to FIGS. 1 to 21.

[0431]

<7. Sixth embodiment>

[Other examples of implementation]

Although the examples of the devices and the system to which the present technology is applied have been described above, the present technology is not limited thereto. For example, the present technology can also be implemented as a processor serving as system large scale integration (LSI) or the like, a module using a plurality of processors or the like, a unit using a plurality of modules or the like, a set in which other functions are further added to the unit, or the like (that is, a configuration of part of the device).

[0432]

plurality of configurations having associated functions as well as the case in which a partial configuration of the multi-functionality is implemented with the development

and manufacturing thereof and the case in which the configuration is implemented to have one function could be seen.

[0434]

A video set 1300 illustrated in FIG. 30 is a multi-functional configuration and serves as a video set obtained by combining a device having a function related to image encoding and decoding (either one or both of which may be given) with a device having other functions associated with the function.

[0435]

As illustrated in FIG. 30, the video set 1300 has a module group of a video module 1311, an external memory 1312, a power management module 1313, a front-end module 1314, and the like and a device having related functions of connectivity 1321, a camera 1322, a sensor 1323, and the like.

[0436]

The module serves as a component with a certain integrated function by integrating some component functions associated with each other. Although a specific physical configuration is optional, for example, the configuration integrated by arranging a plurality of processors having functions, electronic circuit elements such as a resistor and a capacitor, other devices, and the like on a circuit board is possible. In addition, a new module in which the module is combined with other modules or processors is also possible.

[0437]

In the case of the example of FIG. 30, the video module 1311 is a module in

The configuration having the predetermined function may be a logic circuit (hardware configuration) or a CPU, a ROM, a RAM, and the like and a program

(software configuration) to be executed using the CPU, the ROM, the RAM, and the like. For example, the processor may have the logic circuit, the CPU, the ROM, the RAM, and the like, part of the function may be implemented by the logic circuit (hardware configuration), and the other functions may be implemented by the program (software configuration) to be executed in the CPU.

[0439]

An application processor 1331 of FIG. 30 is a processor for executing an application related to image processing. The application to be executed in the application processor 1331 can not only perform a calculation process so as to implement a predetermined function, but also control configurations inside and outside the video module 1311, for example, such as a video processor 1332, if necessary.

[0440]

The video processor 1332 is a processor having a function related to (one or both of) image encoding and decoding.

[0441]

The broadband modem 1333 is a processor (or module) for performing a process related to wireless or wired (or wired and wireless) broadband communication to be performed via a link of a broadband such as the Internet, a public telephone network, or the like. For example, the broadband modem 1333 converts data (digital signal) to be transmitted into an analog signal by performing digital modulation and the like or converts a received analog signal into data (a

RF signal to be transmitted and received via an antenna. For example, the RF module 1334 generates an RF signal by performing frequency conversion and the

like on a baseband signal generated by the broadband modem 1333. In addition, for example, the RF module 1334 generates a baseband signal by performing frequency conversion and the like on an RF signal received via the front-end module 1314.

[0443]

Also, as indicated by a dotted line 1341 in FIG. 30, one processor may be configured by integrating the application processor 1331 and the video processor 1332.

[0444]

The external memory 1312 is a module provided outside the video module 1311 and having a storage device to be used by the video module 1311. Although the storage device of the external memory 1312 may be configured to be implemented by any physical configuration, for example, it is desirable to implement the storage device using a relatively cheap and large-capacity semiconductor memory, for example, such as a dynamic random access memory (DRAM), because the storage device is generally used to store a large volume of data such as image data of frame units in many cases.

[0445]

The power management module 1313 manages and controls power supply to the video module 1311 (each configuration within the video module 1311).

[0446]

The front-end module 1314 is a module for providing a front-end function (a circuit of a transmission/reception end of an antenna side) to the RF module 1334.

filter 1352 performs a filtering process and the like on the RF signal received via the antenna section 1351 and supplies the processed RF signal to the RF module 1334.

The amplification section 1353 amplifies the RF signal supplied from the RF module 1334 and supplies the amplified RF signal to the antenna section 1351.

[0448]

The connectivity 1321 is a module having a function related to a connection to the outside. A physical configuration of the connectivity 1321 is optional. For example, the connectivity 1321 has a configuration having a communication function other than a communication standard corresponding to the broadband modem 1333, external input and output ports, and the like.

[0449]

For example, the connectivity 1321 may be configured to have a module having a communication function based on a wireless communication standard such as Bluetooth (registered trademark), IEEE 802.11 (for example, Wi-Fi (registered trademark)), near field communication (NFC), or infrared data association (IrDA), an antenna for transmitting and receiving a signal based on the standard. In addition, the connectivity 1321 may be configured to have a module with a communication function based on a wired communication standard such as a universal serial bus (USB) or a high-definition multimedia I/F (HDMI) (registered trademark) and a port based on the standard. Further, for example, the connectivity 1321 may be configured to have other data (signal) transmission functions of antenna input and output ports and the like.

[0450]

Also, the connectivity 1321 may be configured to include a device of a

[0451]

The camera 1322 is a module having a function of capturing of an image of

a subject and obtaining image data of the subject. The image data obtained by the image capturing of the camera 1322, for example, is supplied to the video processor 1332 and encoded.

[0452]

The sensor 1323, for example, is a module having an arbitrary sensor function such as a sound sensor, an ultrasonic sensor, an optical sensor, an illuminance sensor, an infrared sensor, an image sensor, a rotation sensor, an angle sensor, an angular velocity sensor, a velocity sensor, an acceleration sensor, an inclination sensor, a magnetic identification sensor, an impact sensor, or a temperature sensor. Data detected by the sensor 1323, for example, is supplied to the application processor 1331 and used by an application or the like.

[0453]

The configuration described above as the module may be implemented as a processor and conversely the configuration described as the processor may be implemented as a module.

[0454]

In the video set 1300 of the configured as described above, it is possible to apply the present technology to the video processor 1332 as will be described later. Therefore, the video set 1300 can be executed as a set to which the present technology is applied.

[0455]

[Configuration example of video processor]

video signal and an audio signal.

[0457]

As illustrated in FIG. 31, the video processor 1332 has a video input processing section 1401, a first image enlargement/reduction section 1402, a second image enlargement/reduction section 1403, a video output processing section 1404, a frame memory 1405, and a memory control section 1406. In addition, the video processor 1332 has an encoding and decoding engine 1407, video elementary stream (ES) buffers 1408A and 1408B, and audio elementary stream (ES) buffers 1409A and 1409B. Further, the video processor 1332 has an audio encoder 1410, an audio decoder 1411, a multiplexer (MUX) 1412, an inverse multiplexing section (demultiplexer (DMUX)) 1413, and a stream buffer 1414.

[0458]

The video input processing section 1401, for example, acquires a video signal input from the connectivity 1321 (FIG. 30) and converts the video signal into digital image data. The first image enlargement/reduction section 1402 performs a format conversion process, an image enlargement/reduction process, or the like on image data. The second image enlargement/reduction section 1403 performs the image enlargement/reduction process on the image data according to a format in a destination to which the image data is output via the video output processing section 1404 or perform format conversion, an image enlargement/reduction process or the like, and the like as in the first image enlargement/reduction section 1420. The video output processing section 1404 performs the format conversion, the conversion into an analog signal, or the like on the image data and outputs a conversion result as a reproduced video signal, for example, to the connectivity 1321 (FIG. 30) and the

DRAM.

[0460]

The memory control section 1406 receives a synchronization signal from the encoding and decoding engine 1407 and controls access of writing/reading to/from the frame memory 1405 according to an access schedule for the frame memory written to an access management table 1406A. The access management table 1406A is updated by the memory control section 1406 according to a process to be executed by the encoding and decoding engine 1407, the first image enlargement/reduction section 1402, the second image enlargement/reduction section 1403, or the like.

[0461]

The encoding and decoding engine 1407 performs an image data encoding process and a process of decoding a video stream which is data in which image data is encoded. For example, the encoding and decoding engine 1407 encodes image data read from the frame memory 1405 and sequentially writes the encoded image data as a video stream to the video ES buffer 1408. In addition, for example, a video stream from the video ES buffer 1408B is sequentially read and decoded and the decoded video stream is sequentially written as image data to the frame memory 1405. The encoding and decoding engine 1407 uses the frame memory 1405 as a work area in the encoding or decoding of the image data. In addition, the encoding and decoding engine 1407, for example, outputs a synchronization signal to the memory control section 1406 at the timing at which processing for every macro block starts.

[0462]

encoder 1410 and supplies the audio stream to the multiplexing section (MUX) 1412. The audio ES buffer 1409B buffers an audio stream supplied from the inverse

multiplexing section (DMUX) 1413 and supplies the audio stream to the audio decoder 1411.

[0464]

The audio encoder 1410, for example, digitally converts an audio signal input from the connectivity 1321 (FIG. 30) or the like, and for example, encodes the digitally converted audio signal in a predetermined scheme such as an MPEG audio scheme or an audio code number 3 (AC3) scheme. The audio encoder 1410 sequentially writes an audio data stream which is data in which an audio signal is encoded to the audio ES buffer 1409A. The audio decoder 1411 decodes an audio stream supplied from the audio ES buffer 1409B and supplies the decoded audio stream as a reproduced audio signal, for example, to the connectivity 1321 (FIG. 30) or the like, for example, by performing the conversion into the analog signal or the like.

[0465]

The multiplexing section (MUX) 1412 multiplexes a video stream and an audio stream. This multiplexing method (that is, a format of the bitstream generated by multiplexing) is optional. In addition, at the time of multiplexing, the multiplexing section (MUX) 1412 can add predetermined header information or the like to the bitstream. That is, the multiplexing section (MUX) 1412 can convert a format of the stream by the multiplexing. For example, the multiplexing section (MUX) 1412 performs conversion into a transport stream which is a bitstream of a transmission format by multiplexing the video stream and the audio stream. In

audio stream from the bitstream read from the stream buffer 1414 (demultiplexes the video stream and the audio stream). That is, the inverse multiplexing section

(DMUX) 1413 can convert the format of the inversely multiplexed stream (inverse conversion to the conversion by the multiplexing section (MUX) 1412). For example, the inverse multiplexing section (DMUX) 1413, for example, acquires a transport stream supplied from the connectivity 1321, the broadband modem 1333, or the like (all in FIG. 30) via the stream buffer 1414 and demultiplexes the acquired transport stream, thereby converting the transport stream into the video stream and the audio stream. In addition, for example, the inverse multiplexing section (DMUX) 1413 can acquire file data read from various types of recording media, for example, by the connectivity 1321 (FIG. 30) via the stream buffer 1414 and inversely multiplex the acquired file data to perform conversion into a video stream and an audio stream.

[0467]

The stream buffer 1414 buffers the bitstream. For example, the stream buffer 1414 buffers the transport stream supplied from the multiplexing section (MUX) 1412 and supplies the transport stream, for example, to the connectivity 1321, the broadband modem 1333, or the like (all in FIG. 30) at a predetermined timing or based on an external request or the like.

[0468]

In addition, for example, the stream buffer 1414 buffers file data supplied from the multiplexing section (MUX) 1412 and supplies the buffer file data, for example, to the connectivity 1321 (FIG. 30) or the like, at a predetermined timing or based on an external request or the like to cause various types of recording media to

In addition, the stream buffer 1414, for example, buffers file data read from various types of recording media in the connectivity 1321 (FIG. 30) and the like and

supplies the file data to the inverse multiplexing section (DMUX) 1413 at a predetermined timing or based on an external request or the like.

[0471]

Next, an example of an operation of the video processor 1332 of this configuration will be described. For example, a video signal input to the video processor 1332 from the connectivity 1321 (FIG. 30) or the like is converted into digital image data of a predetermined scheme such as a 4:2:2Y/Cb/Cr scheme in the video input processing section 1401 and the digital image data is sequentially written to the frame memory 1405. This digital image data is read to the first image enlargement/reduction section 1402 or the second image enlargement/reduction section 1403, the format conversion into a predetermined scheme such as a 4:2:0Y/Cb/Cr scheme and the enlargement/reduction process are performed and the image data is written to the frame memory 1405 again. The image data is encoded by the encoding and decoding engine 1407 and the encoded image data is written as a video stream to the video ES buffer 1408A.

[0472]

In addition, an audio signal input from the connectivity 1321 (FIG. 30) or the like to the video processor 1332 is encoded by the audio encoder 1410 and the encoded audio signal is written as an audio stream to the audio ES buffer 1409A.

[0473]

The video stream of the video ES buffer 1408A and the audio stream of the audio ES buffer 1409A are read to the multiplexing section (MUX) 1412 and

types of recording media.

[0474]

In addition, for example, after the transport stream input from the external network to the video processor 1332 via (any one (FIG. 30) of) the connectivity 1321, the broadband modem 1333, and the like is buffered in the stream buffer 1414, the transport stream is inversely multiplexed by the inverse multiplexing section (DMUX) 1413. In addition, for example, after file data read from various types of recording media in the connectivity 1321 (FIG. 30) and the like and input to the video processor 1332 is buffered in the stream buffer 1414, the file data is inversely multiplexed by the inverse multiplexing section (DMUX) 1413. That is, the transport stream or the file data input to the video processor 1332 is demultiplexed into a video stream and an audio stream by the inverse multiplexing section (DMUX) 1413.

[0475]

An audio signal is reproduced by supplying the audio stream to the audio decoder 1411 via the audio ES buffer 1409B and decoding the audio stream. In addition, after the video stream is written to the video ES buffer 1408B, the video stream is sequentially read and decoded by the encoding and decoding engine 1407 and written to the frame memory 1405. The second image enlargement/reduction section 1403 performs the decoded image data on the enlargement/reduction process and the processed data is written to the frame memory 1405. Then, a video signal is reproduced and output by reading the decoded image data to the video output processing section 1404, converting the format of the decoded image data in a predetermined scheme such as 4:2:2Y/Cb/Cr scheme, and further converting the

according to the first embodiment and the image decoding device (FIG. 14) according to the second embodiment. As described above, the video processor

1332 can obtain effects similar to those described above with reference to FIGS. 1 to 21.

[0477]

Also, in the encoding and decoding engine 1407, the present technology (that is, the functions of the image encoding device and the image decoding device according to the embodiments described above) may be implemented by hardware such as a logic circuit, implemented by software such as an embedded program, or implemented by both thereof.

[0478]

[Other configuration example of video processor]

FIG. 32 illustrates another example of a schematic configuration of the video processor 1332 (FIG. 30) to which the present technology is applied. In the case of the example of FIG. 32, the video processor 1332 has a function of encoding and decoding video data in a predetermined scheme.

[0479]

More specifically, as illustrated in FIG. 32, the video processor 1332 has a control section 1511, a display I/F 1512, a display engine 1513, an image processing engine 1514, and an internal memory 1515. In addition, the video processor 1332 has a codec engine 1516, a memory I/F 1517, a multiplexing/inverse multiplexing section (MUX/DMUX) 1518, a network I/F 1519, and a video I/F 1520.

[0480]

The control section 1511 controls operations of processing sections within

to the program or the like and supplies the control signal to each processing section (that is, controls the operation of each processing section). The sub-CPU 1532

plays an auxiliary role of the main CPU 1531. For example, the sub-CPU 1532 executes a child process, a sub-routine, or the like of a program or the like to be executed by the main CPU 1531. The system controller 1533 controls operations of the main CPU 1531 and the sub-CPU 1532 such as designations of programs to be executed by the main CPU 153 and the sub CPU 1532.

[0482]

The display I/F 1512 outputs image data, for example, to the connectivity 1321 (FIG. 30) and the like, under control of the control section 1511. For example, the display I/F 1512 converts image data of digital data into an analog signal and outputs the analog signal as the reproduced video signal, or the image data of the digital data without change, to the monitor device or the like of the connectivity 1321 (FIG. 30).

[0483]

The display engine 1513 performs various types of conversion processes such as format conversion, size conversion, and color gamut conversion so that image data is suitable for hardware specs of a monitor device and the like for displaying its image under control of the control section 1511.

[0484]

The image processing engine 1514 performs predetermined image processing, for example, such as filtering processing for improving image quality, on the image data under control of the control section 1511.

[0485]

image processing engine 1514, or the codec engine 1516, if necessary (for example, according to a request). Although the internal memory 1515 may be implemented

by any storage device, it is desirable to implement the internal memory 1515 by a semiconductor memory having a relatively small capacity and high response speed (as compared to those of the external memory 1312), for example, such as a static random access memory (SRAM), because a small volume of data such as image data of a block unit or parameters are generally used in many cases.

[0486]

The codec engine 1516 performs a process related to encoding and decoding of image data. An encoding and decoding scheme corresponding to this codec engine 1516 is optional and the number of encoding and decoding schemes may be one or more. For example, the codec engine 1516 may have a codec function of a plurality of encoding and decoding schemes, and the encoding of image data or the decoding of encoded data may be performed in one selected from among the encoding and decoding schemes.

[0487]

In the example illustrated in FIG. 32, the codec engine 1516, for example, has an MPEG-2 video 1541, an AVC/H.264 1542, an HEVC/H.265 1543, an HEVC/H.265 (scalable) 1544, an HEVC/H.265 (multi-view) 1545, and an MPEG-DASH 1551 as functional blocks for processing related to a codec.

[0488]

The MPEG-2 video 1541 is a functional block for encoding or decoding image data in an MPEG-2 scheme. The AVC/H.264 1542 is a functional block for encoding or decoding image data in an AVC scheme. The HEVC/H.265 1543 is a

image data in an MPEG-DASH scheme. The MPEG-DASH is technology for performing video streaming using an HTTP and has one feature that appropriate

encoded data is selected from a plurality of pieces of encoded data in which prepared resolutions and the like are different from each other in units of segments and the selected encoded data is transmitted. The MPEG-DASH 1551 performs generation of a stream based on a standard, transmission control of the stream, and the like and uses the MPEG-2 video 1541 to the HEVC/H.265 (multi-view) 1545 described above for encoding and decoding of the image data.

[0490]

The memory I/F 1517 is an I/F for the external memory 1312. Data supplied from the image processing engine 1514 or the codec engine 1516 is supplied to the external memory 1312 via the memory I/F 1517. In addition, data read from the external memory 1312 is supplied to the video processor 1332 (the image processing engine 1514 or the codec engine 1516) via the memory I/F 1517.

[0491]

The multiplexing/inverse multiplexing section (MUX/DMUX) 1518 performs multiplexing or demultiplexing of various types of data related to an image such as an encoded data bitstream, image data, and a video signal. A method of this multiplexing/demultiplexing is optional. For example, at the time of multiplexing, the multiplexing/inverse multiplexing section (MUX/DMUX) 1518 can not only integrate a plurality of piece data into one piece, but also add predetermined header information or the like to the data. In addition, at the time of inverse multiplexing, the multiplexing/inverse multiplexing section (MUX/DMUX) 1518 can not only divide one piece of data into a plurality of pieces of data, but also add predetermined

[0492]

The network I/F 1519, for example, is an I/F for the broadband modem 1333

or the connectivity 1321 (both in FIG. 30) or the like. The video I/F 1520, for example, is an I/F for the connectivity 1321 or the camera 1322 (both in FIG. 30) or the like.

[0493]

Next, an example of an operation of this video processor 1332 will be described. For example, when a transport stream is received from the external network via the connectivity 1321 or the broadband modem 1333 (both in FIG. 30) or the like, the transport stream is supplied to the multiplexing/inverse multiplexing section (MUX/DMUX) 1518 via the network I/F 1519 and inversely multiplexed, and the transport stream is decoded by the codec engine 1516. For example, the image processing engine 1514 performs predetermined image processing on image data obtained by the decoding of the codec engine 1516, the display engine 1513 performs predetermined conversion on the processed image data, the converted image data, for example, is supplied to the connectivity 1321 (FIG. 30) or the like via the display I/F 1512, and its image is displayed on a monitor. In addition, for example, the codec engine 1516 re-encodes image data obtained by the decoding of the codec engine 1516, the multiplexing/inverse multiplexing section (MUX/DMUX) 1518 multiplexes the re-encoded image data to perform conversion into file data, the file data is output, for example, to the connectivity 1321 (FIG. 30) or the like, via the video I/F 1520, and the output file data is recorded to various types of recording media.

[0494]

the processed image data, the converted image data, for example, is supplied to the connectivity 1321 (FIG. 30) or the like via the display I/F 1512, and its image is

displayed on a monitor. In addition, for example, the codec engine 1516 re-encodes image data obtained by the decoding of the codec engine 1516, the multiplexing/inverse multiplexing section (MUX/DMUX) 1518 multiplexes the re-encoded image data to perform conversion into a transport stream, the transport stream is supplied, for example, to the connectivity 1321 or the broadband modem 1333 (both in FIG. 30) or the like, via the network I/F 1519 and transmitted to another device (not illustrated).

[0495]

Also, the exchange of image data between processing sections within the video processor 1332 or other data, for example, is performed using the internal memory 1515 or the external memory 1312. In addition, the power management module 1313, for example, controls power supply to the control section 1511.

[0496]

When the present technology is applied to the video processor 1332 configured as described above, it is only necessary to apply the present technology according to each embodiment described above to the codec engine 1516. That is, for example, it is only necessary to configure the codec engine 1516 having functional blocks that achieve the image encoding device (FIG. 4) according to the first embodiment and the image decoding device (FIG. 14) according to the second embodiment. As described above, the video processor 1332 can obtain effects similar to those described above with reference to FIGS. 1 to 21.

[0497]

have been shown above, the configuration of the video processor 1332 is optional and may be different from the above-described two examples. In addition, although

the video processor 1332 is configured as one semiconductor chip, the video processor 1332 may be configured as a plurality of semiconductor chips. For example, the video processor 1332 may be configured to be three-dimensional stacked LSI in which a plurality of semiconductors are stacked. In addition, the video processor 1332 may be configured to be implemented by a plurality of pieces of LSI.

[0499]

[Application example for device]

The video set 1300 can be embedded in various types of devices for processing image data. For example, the video set 1300 can be embedded in the television device 900 (FIG. 23), the mobile phone 920 (FIG. 24), the recording/reproduction device 940 (FIG. 25), the image capturing device 960 (FIG. 26), or the like. By embedding the video set 1300, its device can obtain effects similar to those described above with reference to FIGS. 1 to 21.

[0500]

In addition, the video set 1300, for example, can also be embedded in a terminal device such as the PC 1004, the AV device 1005, the tablet device 1006, or the mobile phone 1007 in the data transmission system 1000 of FIG. 27, the broadcasting station 1101 and the terminal device 1102 in the data transmission system 1100 of FIG. 28, the image capturing device 1201 and the scalable coded data storage device 1202 in the image capturing system 1200 of FIG. 29, and the like. By embedding the video set 1300, the device can obtain effects similar to those

1332 can be executed as the video processor to which the present technology is applied. In addition, for example, the processor indicated by the dotted line 1341 as

described above, the video module 1311, and the like can be executed as the processor, the module, or the like to which the present technology is applied. Further, for example, the video module 1311, the external memory 1312, the power management module 1313, and the front-end module 1314 can be combined and executed as a video unit 1361 to which the present technology is applied. Any configuration can obtain effects similar to those described above with reference to FIGS. 1 to 21.

[0502]

That is, any configuration including the video processor 1332 can be embedded in various types of devices for processing image data as in the video set 1300. For example, the video processor 1332, the processor indicated by the dotted line 1341, the video module 1311, or the video unit 1361 can be embedded in the television device 900 (FIG. 23), the mobile phone 920 (FIG. 24), the recording/reproduction device 940 (FIG. 25), the image capturing device 960 (FIG. 26), the terminal device such as the PC 1004, the AV device 1005, the tablet device 1006, or the mobile phone 1007 in the data transmission system 1000 of FIG. 27, the broadcasting station 1101 and the terminal device 1102 in the data transmission system 1100 of FIG. 28, the image capturing device 1201 and the scalable coded data storage device 1202 in the image capturing system 1200 of FIG. 29, and the like. Further, the video processor 1332 can be embedded in each device of the content reproduction system of FIG. 33or the wireless communication system of FIG. 39. By embedding any configuration to which the present technology is applied, its

prepared resolutions and the like are different from each other in units of segments.

[0504]

<8. Application example of MPEG-DASH>

[Outline of content reproduction system]

First, with reference to FIGS. 33 to 35, the content reproduction system to which the present technology is applicable will be schematically described.

[0505]

Hereinafter, first, a basic configuration common in these embodiments will be described with reference to FIGS. 33 and 34.

[0506]

FIG. 33 is an explanatory diagram illustrating a configuration of a content reproduction system. As illustrated in FIG. 33, the content reproduction system includes content servers 1610 and 1611, a network 1612, and a content reproduction device 1620 (client device).

[0507]

The content servers 1610 and 1611 and the content reproduction device 1620 are connected via the network 1612. The network 1612 is a wired or wireless transmission path of information to be transmitted from a device connected to the network 1612.

[0508]

For example, the network 1612 may include public line networks such as the Internet, a telephone line network, and a satellite communication network or various types of local area networks (LANs), wide area networks (WANs), and the like including Ethernet (registered trademark). In addition, the network 1612 may

"moov."

[0510]

In addition, the content data may music data such as music, lectures, and a radio programs, video data such as movies, television programs, video programs, photographs, documents, paintings, and charts, games, software, and the like.

[0511]

Here, for a reproduction request of content from the content reproduction device 1620, the content server 1610 generates a plurality of data files at different bit rates in relation to the same content. In addition, for a reproduction request of content from the content reproduction device 1620, the content server 1611 transmits information of a uniform resource locator (URL) of the content server 1610 to the content reproduction device 1620 by including parameter information added to the URL in the content reproduction device 1620 in the URL information of the content server 1610. Hereinafter, with reference to FIG. 34, relevant items will be specifically described.

[0512]

FIG. 32 is an explanatory diagram illustrating a flow of data in the content reproduction system of FIG. 33. The content server 1610 encodes the same content data at different bit rates and generates, for example, a file A of 2 Mbps, a file B of 1.5 Mbps, and a file C of 1 Mbps as illustrated in FIG. 34. Relatively, the file A has a high bit rate, the file B has a standard bit rate, and the file C has a low bit rate.

[0513]

In addition, as illustrated in FIG. 34, encoded data of each file is divided into a plurality of segments. For example, the encoded data of the file A is divided into

refresh (IDR)-picture in video encoding of AVC/H.264). For example, when video data of 30 frames per sec is encoded in a GOP of a fixed length of 15 frames, each

segment may be video and audio encoded data of 2 sec corresponding to 4 GOPs or video and audio encoded data of 10 sec corresponding to 20 GOPs.

[0515]

In addition, reproduction ranges (ranges of time positions from the beginning of content) by segments having the same placement order in each file are the same. For example, when reproduction ranges of the segment "A2," the segment "B2," and the segment "C2" are the same and each segment is encoded data of 2 sec, all reproduction ranges of the segment "A2," the segment "B2," and the segment "C2" are 2 to 4 sec of content.

[0516]

When the files A to C configured from the above-described plurality of segments are generated, the content server 1610 stores the files A to C. Then, as illustrated in FIG. 34, the content server 1610 sequentially transmits segments constituting different files to the content reproduction device 1620, and the content reproduction device 1620 reproduces the received segments according to streaming reproduction.

[0517]

Here, the content server 1610 according to this embodiment transmits a play list file (hereinafter referred to as MPD) including bit rate information and access information of encoded data to the content reproduction device 1620, and the content reproduction device 1620 selects any one bit rate of a plurality of bit rates based on the MPD and requests the content server 1610 to transmit segments corresponding to plurality of pieces of encoded having different bit rates (bandwidths). For example, the MPD illustrated in FIG. 35 indicates that there are encoded data of 256 Kbps,

1.024 Mbps, 1.384 Mbps, 1.536 Mbps, and 2.048 Mbps, and includes access information related to encoded data. The content reproduction device 1620 can dynamically change the bit rate of encoded data to be reproduced according to streaming reproduction based on this MPD.

[0520]

Also, although a portable terminal is illustrated as an example of the content reproduction device 1620 in FIG. 33, the content reproduction device 1620 is not limited to this example. For example, the content reproduction device 1620 may be an information processing device such as a PC, a home video processing device (a digital versatile disc (DVD) recorder, a video cassette recorder, or the like), personal digital assistants (PDA), a home game device, or a home appliance device. In addition, the content reproduction device 1620 may be an information processing device such as a mobile phone, a personal handy-phone system (PHS), a portable music reproduction device, a portable video processing device, or a portable game device.

[0521]

[Configuration of content server 1610]

With reference to FIGS. 33 to 35, the outline of the content reproduction system has been described. Next, with reference to FIG. 36, the configuration of the content server 1610 will be described.

[0522]

FIG. 36 is a functional block diagram illustrating the configuration of the

section 1631 generates the MPD as illustrated in FIG. 35 when encoded data of 256 Kbps, 1.024 Mbps, 1.384 Mbps, 1.536 Mbps, and 2.048 Mbps is generated.

[0524]

The storage section 1632 stores a plurality of pieces of encoded data having different bit rates and the MPD generated by the file generation section 1631. The storage section 1632 may be a storage medium such as a non-volatile memory, a magnetic disk, an optical disk, an magneto optical (MO) disk, or the like. As the non-volatile memory, for example, an electrically erasable programmable read-only memory (EEPROM) and an erasable programmable ROM (EPROM) may be cited. Also, as the magnetic disk, a hard disk, a discoid magnetic disk, and the like may be cited. Further, as the optical disk, a compact disc (CD), a DVD-recordable (DVD-R), a Blu-ray disc (BD) (registered trademark)), and the like may be cited.

[0525]

The communication section 1633 is an I/F with the content reproduction device 1620, and communicates with the content reproduction device 1620 via the network 1612. In further detail, the communication section 1633 has a function serving as an HTTP server which communicates with the content reproduction device 1620 according to an HTTP. For example, the communication section 1633 transmits an MPD to the content reproduction device 1620, extracts encoded data requested based on the MPD from the content reproduction device 1620 according to the HTTP, and transmits the encoded data to the content reproduction device 1620 as an HTTP response.

[0526]

[Configuration of content reproduction device 1620]

1652, a reproduction section 1653, a selection section 1654, and a current location acquisition section 1656.

[0528]

The communication section 1651 is an I/F with the content server 1610, requests the content server 1610 to provide data, and acquires data from the content server 1610. In further detail, the communication section 1651 has a function serving as an HTTP client for communicating with the content reproduction device 1620 according to the HTTP. For example, the communication section 1651 can selectively acquire an MPD or segments of encoded data from the content server 1610 by employing an HTTP range.

[0529]

The storage section 1652 stores various types of information related to reproduction of content. For example, the storage section 1652 sequentially buffers segments acquired by the communication section 1651 from the content server 1610. The segments of the encoded data buffered in the storage section 1652 are sequentially supplied to the reproduction section 1653 in first in first output (FIFO).

[0530]

In addition, the storage section 1652 stores a definition for accessing a URL to which the communication section 1651 adds a parameter based on an addition instruction of the parameter for the URL of the content described in the MPD requested from the content sever 1611 to be described later.

[0531]

The reproduction section 1653 sequentially reproduces segments supplied from the storage section 1652. Specifically, the reproduction section 1653 performs

segments "A1," "B2," and "A3" from the content server 1610 as illustrated in FIG. 34.

[0533]

The current location acquisition section 1656 may be a section for acquiring a current location of the content reproduction device 1620, and, for example, may be constituted of a module for acquiring a current location of a global positioning system (GPS) receiver or the like. In addition, the current location acquisition section 1656 may be a section for acquiring a current location of the content reproduction device 1620 using a wireless network.

[0534]

[Configuration of content server 1611]

FIG. 38 is an explanatory diagram illustrating a configuration example of the content server 1611. As illustrated in FIG. 38, the content server 1611 includes a storage section 1671 and a communication section 1672.

[0535]

The storage section 1671 stores information of a URL of an MPD. The information of the URL of the MPD is transmitted from the content server 1611 to the content reproduction device 1620 according to a request from the content reproduction device 1620 for requesting reproduction of content. In addition, when the information of the URL of the MPD for the content reproduction device 1620 is provided, the storage section 1671 stores definition information when the content reproduction device 1620 adds a parameter to the URL written in the MPD.

[0536]

The communication section 1672 is an I/F with the content reproduction

[0537]

For the parameter to be added to the URL of the MPD in the content

reproduction device 1620, it is possible to set various parameters in definition information to be shared by the content server 1611 and the content reproduction device 1620. For example, information of a current location of the content reproduction device 1620, a user ID of a user using the content reproduction device 1620, a memory size of the content reproduction device 1620, a capacity of a storage of the content reproduction device 1620, and the like can be added to the URL of the MPD in the content reproduction device 1620.

[0538]

In the content reproduction system of the above-described configuration, it is possible to obtain effects similar to those described above with reference to FIGS. 1 to 21 by applying the present technology as described above with reference to FIGS. 1 to 21.

[0539]

That is, the encoder 1641 of the content server 1610 has a function of the image encoding device (FIG. 4) according to the above-described embodiment. In addition, the reproduction section 1653 of the content reproduction device 1620 has a function of the image decoding device (FIG. 14) according to the above-described embodiment. Thereby, it is possible to improve the encoding efficiency of encoding or decoding of an MV in a multi-view image.

[0540]

In addition, because it is possible to limit the V direction of the inter-view MV by transmitting and receiving encoded data according to the present technology

wireless communication system to which the present technology is applicable will be described.

[0542]

Initially, radio packet transmission and reception until a specific application is operated by setting up a peer to peer (P2P) connection are done.

[0543]

Next, before a connection in a second layer is made, radio packet transmission and reception until the P2P connection is set up and the specific application is operated after the specific application is designated are done. Thereafter, after the connection in the second layer, radio packet transmission and reception when the specific application is activated are done.

[0544]

[Communication example when specific application operation starts]

FIGS. 39 and 40 are examples of radio packet transmission and reception until a specific application is operated by setting up the above-described P2P connection and are sequence charts illustrating a communication processing example by each device based on wireless communication. Specifically, an example of a setup procedure of a direct connection leading to a connection in a Wi-Fi direct standard (also referred to as Wi-Fi P2P) standardized in Wi-Fi alliance is shown.

[0545]

Here, in the Wi-Fi direct, a plurality of wireless communication devices detects the existence of one another (device discovery and service discovery). Then, a direct connection is set up by performing device authentication in Wi-Fi protected setup (WPS) with a selected device when connection device selection is performed.

packet exchange is necessary even in exchange of Authentication Request/Response or the like. However, in FIGS. 39 and 40, the illustration of these packet exchanges

is omitted and only a connection from the second time is shown.

[0547]

Also, although a communication process example between a first wireless communication device 1701 and a second wireless communication device 1702 is illustrated in FIGS. 39 and 40, the same is true for communication processes with other wireless communication devices.

[0548]

Initially, the device discovery is performed between the first wireless communication device 1701 and the second wireless communication device 1702 (1711). For example, the first wireless communication device 1701 transmits a probe request (response request signal), and receives a probe response (response signal) for the probe request from the second wireless communication device 1702. Thereby, the first wireless communication device 1701 and the second wireless communication device 1702 can discover the location of each other. In addition, it is possible to acquire a device name or type (TV, PC, smartphone, or the like) of a partner by the device discovery.

[0549]

Subsequently, the service discovery is performed between the first wireless communication device 1701 and the second wireless communication device 1702 (1712). For example, the first wireless communication device 1701 transmits a service discovery query for querying a service corresponding to the second wireless communication device 1702 discovered in the device discovery. Then, the first

Subsequently, the user performs an operation (connection partner selection operation) of selecting a connection partner (1713). This connection partner

selection operation occurs in only any one of the first wireless communication device 1701 and the second wireless communication device 1702. For example, a connection partner selection picture is displayed on the display section of the first wireless communication device 1701, and the second wireless communication device 1702 is selected as the connection partner in the connection partner selection picture according to the user operation.

[0551]

When the user performs the connection partner selection operation (1713), group owner negotiation is performed between the first wireless communication device 1701 and the second wireless communication device 1702 (1714). In FIGS. 39 and 40, an example in which the first wireless communication device 1701 becomes a group owner 1715 and the second wireless communication device 1702 becomes a client 1716 according to a result of the group owner negotiation is illustrated.

[0552]

Subsequently, processes 1717 to 1720 are performed between the first wireless communication device 1701 and the second wireless communication device 1702, so that a direct connection is set up. That is, association (L2 (second layer) link setup) 1717 and secure link setup 1718 are sequentially performed. In addition, IP address assignment 1719 and L4 setup 1720 on L3 by a simple service discovery protocol (SSD) or the like are sequentially performed. Also, L2 (layer 2) refers to a second layer (data link layer), L3 (layer 3) refers to a third layer (network layer), and

operation picture is displayed on the display section of the first wireless communication device 1701, and the specific application is selected by the user

operation in this application designation/activation operation picture.

[0554]

When the user performs the application designation/activation operation (1721), the specific application corresponding to this application designation/activation operation is executed between the first wireless communication device 1701 and the second wireless communication device 1702 (1722).

[0555]

Here, the case in which a connection between an access point (AP) and a station (STA) is set up within a range of specs (specs standardized in IEEE 802.11) before the Wi-Fi direct standard is assumed. In this case, before a connection is made in the second layer (before association in the terminology of IEEE 802.11), it is difficult to know what device is connected in advance.

[0556]

On the other hand, as illustrated in FIGS. 39 and 40, in the Wi-Fi direct, it is possible to acquire information of a connection partner when a connection candidate partner is found in the device discovery or service discovery (option). The information of the connection partner, for example, is a basic type of device, a corresponding specific application, or the like. Then, it is possible to allow the user to select a connection partner based on the acquired information of the connection partner.

[0557]

[0558]

[Configuration example of frame format]

FIG. 41 is a diagram schematically illustrating a configuration example of a format of a frame to be transmitted and received in a communication process by each device used as the basis of the present technology. That is, in FIG. 41, a configuration example of a medium access control (MAC) frame for setting up the connection in the second layer is illustrated. Specifically, this is an example of a frame format of an association request/response 1787 for implementing the sequence illustrated in FIG. 42.

[0559]

Also, fields from frame control 1751 to sequence control 1756 are a MAC header. In addition, when an association request is transmitted, B3B2 = "0b00" and B7B6B5B4 = "0b0000" are set in the frame control 1751. In addition, when an association response is encapsulated, B3B2 = "0b00" and B7B6B5B4 = "0b0001" are set in the frame control 1751. Also, "0b00" represents "00" in the binary notation, "0b0000" represents "0000" in the binary notation, and "0b0001" represents "0001" in the binary notation.

[0560]

Here, the MAC frame illustrated in FIG. 41 is basically an association request/response frame format written in Sections 7.2.3.4 and 7.2.3.5 of IEEE 802.11-2007 specification document. However, the format is different in that an independently extended IE as well as an information element (hereinafter abbreviated as an IE) is included.

[0561]

indicating a type of vendor specific IE is first provided. Then, thereafter, a plurality of subelements 1766 is configured to be stored.

[0563]

As the content of the sub-element 1766, a name 1767 of a specific application to be used or a device role 1768 when the specific application operates is considered to be included. In addition, the specific application, information (information for L4 setup) 1769 such as a port number to be used for control, and information (capability information) related to capability is considered to be included. Here, for example, when a designated specific application is DLNA, the capability information is information for specifying correspondence to audio transmission/reproduction, correspondence to video transmission/reproduction, and the like.

[0564]

In the wireless communication system of the above-described configuration, it is possible to obtain an advantageous effect similar to that described above with reference to FIGS. 1 to 21 by applying the present technology described above with reference to FIGS. 1 to 21. That is, it is possible to improve the encoding efficiency of encoding or decoding of an MV in a multi-view image. In addition, in the above-described wireless communication system, it is possible to improve the encoding efficiency of encoding or decoding of an MV in a multi-view image by transmitting and receiving data encoded by the present technology.

[0565]

Also, in the present description, an example in which various types of information such as parameters of a deblocking filter or parameters of an adaptive

linked at the time of decoding. That is, the information may be transmitted on a separate transmission path from an image (or bitstream). In addition, the

information may be recorded on a separate recording medium (or a separate recording area of the same recording medium) from the image (or bitstream). Further, the information and the image (or the bitstream), for example, may be associated with each other in an arbitrary unit such as a plurality of frames, one frame, or a portion within the frame.

[0566]

The preferred embodiments of the present disclosure have been described above with reference to the accompanying drawings, whilst the present invention is not limited to the above examples, of course. A person skilled in the art may find various alterations and modifications within the scope of the appended claims, and it should be understood that they will naturally come under the technical scope of the present disclosure.

[0567]

Additionally, the present technology may also be configured as below.

(1)

An image processing device including:

a predicted vector generation section configured to generate a predicted vector for use in encoding of a motion vector (MV) of a current block by scaling the MV of a reference block, which is a block of a position shifted from a position of the current block in an image of a different view, by a disparity obtained from a periphery of the current block in an image of a non-base view according to a reference destination of the current block and a reference destination of the reference

generation section generates the predicted vector by scaling the MV of the reference block according to a reference image picture order count (POC) of the current block

and a reference image POC of the reference block and employing the scaled MV as a candidate for the predicted vector.

(3)

The image processing device according to (1) or (2), further including:

a transmission section configured to transmit the MV of the current block encoded by the MV encoding section and the encoded stream generated by the encoding section.

(4)

An image processing method including:

generating, by an image processing device, a predicted vector for use in encoding of an MV of a current block by scaling the MV of a reference block, which is a block of a position shifted from a position of the current block in an image of a different view, by a disparity obtained from a periphery of the current block in an image of a non-base view according to a reference destination of the current block and a reference destination of the reference block;

encoding, by the image processing device, the MV of the current block using the generated predicted vector; and

generating, by the image processing device, an encoded stream by encoding the image in units having a hierarchical structure.

(5)

An image processing device including:

a predicted vector generation section configured to generate a predicted

a decoding section configured to generate an image by decoding an encoded stream encoded in units having a hierarchical structure.

(6)

The image processing device according to (5), wherein the predicted vector generation section generates the predicted vector by scaling the MV of the reference block according to a reference image POC of the current block and a reference image POC of the reference block and employing the scaled MV as a candidate for the predicted vector.

(7)

The image processing device according to (5) or (6), further including:

a reception section configured to receive the encoded stream and the encoded MV of the current block.

(8)

An image processing method including:

generating, by an image processing device, a predicted vector for use in encoding of an MV of a current block by scaling the MV of a reference block, which is a block of a position shifted from a position of the current block in an image of a different view, by a disparity obtained from a periphery of the current block in an image of a non-base view according to a reference destination of the current block and a reference destination of the reference block;

decoding, by the image processing device, the MV of the current block using the generated predicted vector; and

generating, by the image processing device, an image by decoding an encoded stream encoded in units having a hierarchical structure.

36-1, 36-N, 36-M MV memory

51 motion prediction mode generation section

52 automatic reference index generation section
53 AMVP mode vector prediction section
54 M/S mode vector prediction section
55 mode determination section
61 vector search section
62 predicted image generation section
63 vector cost determination section
64 spatial MV memory
65, 66 predicted vector generation section
67 switch
68 subtraction section
69 POC transform section
81 predicted vector index generation section
82 intra-view reference vector generation section
83 inter-view reference vector generation section
211-1, 211-N, 211-M decoder
222 lossless decoding section
233-1, 233-N, 233-M DPB
231 motion compensation section
229-1, 229-N, 229-M MV memory
251 automatic reference index generation section
252 AMVP mode vector prediction section

281 intra-view reference vector generation section
282 inter-view reference vector generation section

CLAIMS

Claim 1

An image processing device comprising:

a predicted vector generation section configured to generate a predicted vector for use in encoding of a motion vector (MV) of a current block by scaling the MV of a reference block, which is a block of a position shifted from a position of the current block in an image of a different view, by a disparity obtained from a periphery of the current block in an image of a non-base view according to a reference destination of the current block and a reference destination of the reference block;

an MV encoding section configured to encode the MV of the current block using the predicted vector generated by the predicted vector generation section; and

an encoding section configured to generate an encoded stream by encoding the image in units having a hierarchical structure.

Claim 2

The image processing device according to claim 1, wherein the predicted vector generation section generates the predicted vector by scaling the MV of the reference block according to a reference image picture order count (POC) of the current block and a reference image POC of the reference block and employing the scaled MV as a candidate for the predicted vector.

An image processing method comprising:

generating, by an image processing device, a predicted vector for use in

encoding of an MV of a current block by scaling the MV of a reference block, which is a block of a position shifted from a position of the current block in an image of a different view, by a disparity obtained from a periphery of the current block in an image of a non-base view according to a reference destination of the current block and a reference destination of the reference block;

encoding, by the image processing device, the MV of the current block using the generated predicted vector; and

generating, by the image processing device, an encoded stream by encoding the image in units having a hierarchical structure.

Claim 5

An image processing device comprising:

a predicted vector generation section configured to generate a predicted vector for use in encoding of an MV of a current block by scaling the MV of a reference block, which is a block of a position shifted from a position of the current block in an image of a different view, by a disparity obtained from a periphery of the current block in an image of a non-base view according to a reference destination of the current block and a reference destination of the reference block;

an MV decoding section configured to decode the MV of the current block using the predicted vector generated by the predicted vector generation section; and

a decoding section configured to generate an image by decoding an encoded stream encoded in units having a hierarchical structure.

Claim 7