DESCRIPTION

Title of the Invention:

Decoder, Encoder, Decoding Method, And Encoding Method

Field of the Invention

[0001] The present disclosure relates to a decoder, an encoder, a decoding method, and an encoding method.

Background Art

[0002] WO2023/050431 discloses a determination method for a computing capability resource related to the background art. A control node includes at least one of a terminal side computing capability control node and a network side computing capability control node. The control node receives first information indicating resource information of a computing capability available to the terminal side computing capability control node. The control node determines context information of the computing capability resource based on the first information.

Summary of the Invention

[0004] An object of the present disclosure is to improve execution accuracy of task processing.

[0005] A decoder according to one aspect of the present disclosure includes circuitry and a memory connected to the circuitry, in which the circuitry, in operation, decodes, from a bitstream, an image and an identifier associated with the image, the identifier indicating at least one of (i) the image being suitable for task processing corresponding to the identifier and (ii) the image being not suitable for the task processing.

Brief Description of the Drawings

[0006] Fig. 1 is a view illustrating a simplified configuration of an image processing system according to an embodiment of the present disclosure.

Fig. 2 is a view illustrating a simplified configuration of circuitry included in an encoder.

Fig. 3 is a flowchart showing processing executed by circuitry included in an encoder.

Fig. 4 is a view illustrating a first example of image analysis by a setting unit.

Fig. 5 is a view illustrating a second example of image analysis by a setting unit.

Fig. 6 is a view illustrating a third example of image analysis by a setting unit.

Fig. 7 is a view illustrating image data and identifier data.

Fig. 8 is a view illustrating a first example of syntax regarding setting of an identifier.

Fig. 9 is a view illustrating a first setting example of a value of an identifier in syntax.

Fig. 10 is a view illustrating a second example of syntax regarding setting of an identifier.

Fig. 11 is a view illustrating a second setting example of a value of an identifier in syntax.

Fig. 12 is a view illustrating a simplified configuration of packet data storing each encoded image.

Fig. 13 is a view illustrating a simplified configuration of circuitry included in a decoder.

Fig. 14 is a flowchart showing processing executed by circuitry included in a decoder.

Fig. 15 is a view illustrating a first example of image selection based on identifier data.

Fig. 16 is a view illustrating a second example of image selection based on identifier data.

Fig. 17 is a view illustrating image data and identifier data.

Fig. 18 is a view illustrating an example of syntax regarding setting of an identifier.

Fig. 19 is a view illustrating an example of syntax regarding setting of an identifier.

Fig. 20 is a view illustrating a first setting example of a value of an identifier in syntax, corresponding to Fig. 9.

Fig. 21 is a view illustrating a second setting example of a value of an identifier in syntax, corresponding to Fig. 10.

Fig. 22 is a view illustrating an example of syntax regarding setting of an identifier on the basis of Fig. 18.

Fig. 23 is a view illustrating an example of syntax regarding setting of an identifier on the basis of Fig. 19.

Fig. 24 is a view illustrating an example of syntax regarding setting of task processing.

Fig. 25 is a view illustrating a simplified configuration of circuitry included in an encoder.

Fig. 26 is a view illustrating a simplified configuration of circuitry included in an encoder.

Fig. 27 is a view illustrating a simplified configuration of an image processing system according to a seventh modification.

Fig. 28 is a view illustrating an example of setting information of task processing.

Fig. 29 is a view illustrating a simplified configuration of circuitry included in an encoder.

Fig. 30 is a flowchart showing processing executed by circuitry included in an encoder.

Fig. 31 is a view illustrating image data, first identifier data, and second identifier data.

Fig. 32 is a view illustrating a bit string including first identifier data and second

identifier data.

Fig. 33A is a view illustrating a first example of syntax regarding setting of a first identifier and a second identifier.

Fig. 33B is a view illustrating a first example of syntax regarding setting of a first identifier and a second identifier.

Fig. 34A is a view illustrating a second example of syntax regarding setting of a first identifier and a second identifier.

Fig. 34B is a view illustrating a second example of syntax regarding setting of a first identifier and a second identifier.

Fig. 35 is a view illustrating a simplified configuration of circuitry included in a decoder.

Fig. 36 is a flowchart showing processing executed by circuitry included in a decoder.

Fig. 37 is a view illustrating an example of selection of a first image and a second image based on first identifier data and second identifier data.

Fig. 38 is a view illustrating a setting example of a value of an identifier in syntax.

Fig. 39 is a view illustrating an example of selection of a first image and a second image based on the first identifier data and the second identifier data.

Fig. 40 is a view illustrating an example of syntax regarding setting of task processing.

Fig. 41 is a view illustrating a simplified configuration of a bitstream.

Fig. 42 is a view illustrating an example of syntax regarding setting of a first identifier and a second identifier.

Fig. 43 is a view illustrating an example of syntax regarding setting of a first identifier and a second identifier.

Fig. 44 is a view illustrating an example of syntax regarding setting of a first identifier and a second identifier.

Fig. 45 is a view illustrating an example of syntax regarding setting of a first identifier and a second identifier.

Description of Embodiments

[0007] (Knowledge Underlying Present Disclosure)

An image processing system according to the background art includes an encoder, a decoder, and a task processing unit. The encoder acquires image data of a moving image captured by a camera, and encodes an image included in the acquired image data into a bitstream. The encoder transmits the bitstream to the decoder. The decoder receives a bitstream, decodes an image from the received bitstream, and inputs the decoded image to the task processing unit. The task processing unit executes task processing based on the image input from the decoder.

[0008] The task processing includes human vision and a machine task. Human vision is visual recognition or viewing of a moving image by a human such as an operator or a user. Machine tasks include various types of task processing such as object detection, object tracking, object segmentation, action recognition, or pose estimation using a machine-learned estimation model.

[0009] The processing speed varies depending on the type of task processing depending on the processing capability of the hardware resource that executes each task processing. Therefore, the decoder performs thinning processing on a plurality of images decoded from the bitstream, and inputs the image after the thinning processing to the task processing unit. For example, the decoder thins out images of even frames of decoded 60 fps images to input 30 fps images to the task processing unit.

[0010] However, in a case where an image important for the task processing is included in the thinned out images of even frames, the estimation accuracy of the estimation model decreases, and as a result, the execution accuracy of the task processing decreases.

[0011] In order to solve such problem, the present inventor has found that the above problem can be solved by including, in a bitstream, information indicating whether or not each image is suitable for task processing and transmitting it from an encoder to a decoder, and the decoder executing thinning processing according to the task processing based on the information, and has arrived at the present disclosure.

[0012] Next, each aspect of the present disclosure will be described.

[0013] A decoder according to a first aspect of the present disclosure includes: circuitry; and a memory connected to the circuitry, in which the circuitry, in operation, decodes, from a bitstream, an image and an identifier associated with the image, the identifier indicating at least one of (i) the image being suitable for task processing corresponding to the identifier and (ii) the image being not suitable for the task processing.

[0014] According to the first aspect, the identifier indicates at least one of a fact that the image is suitable for task processing corresponding to the identifier and a fact that the image is not suitable for the task processing. Therefore, at least one image to be used for the task processing can be appropriately selected based on the identifier from among the plurality of images decoded from the bitstream, whereby the execution accuracy of the task processing can be improved.

[0015] In the decoder according to a second aspect of the present disclosure, in the first aspect, the image may include a plurality of images, and the circuitry may further select, based on the identifier, at least one image to be used for the task processing from the plurality of images.

[0016] According to the second aspect, an image suitable for the task processing can be selected based on the identifier, whereby the execution accuracy of the task processing can be improved.

[0017] In the decoder according to a third aspect of the present disclosure, in the first or second

aspect, the task processing may include a machine task.

[0018] According to the third aspect, an image suitable for the machine task can be selected based on the identifier, whereby the machine task can be appropriately executed.

[0019] In the decoder according to a fourth aspect of the present disclosure, in any one of the first to third aspects, the task processing may include human vision.

[0020] According to the fourth aspect, an image suitable for human vision can be selected based on the identifier, whereby human vision can be appropriately executed.

[0021] In the decoder according to a fifth aspect of the present disclosure, in any one of the first to fourth aspects, the identifier may include a plurality of identifiers, the task processing may include multiple types of task processing, and the plurality of identifiers may correspond to different types of task processing among the multiple types of task processing.

[0022] According to the fifth aspect, the plurality of identifiers can correspond to a plurality of pieces of task processing.

[0023] In the decoder according to a sixth aspect of the present disclosure, in the fifth aspect, the plurality of identifiers may include a first identifier and a second identifier, the task processing corresponding to the first identifier may include a machine task, and the task processing corresponding to the second identifier may include human vision.

[0024] According to the sixth aspect, an image suitable for the machine task can be selected based on the first identifier, whereby the machine task can be appropriately executed. An image suitable for human vision can be selected based on the second identifier, whereby human vision can be appropriately executed. That is, it becomes possible to appropriately execute the machine task or the human vision based on the first identifier and the second identifier according to the property of the image.

[0025] In the decoder according to a seventh aspect of the present disclosure, in any one of the first to sixth aspects, the identifier may include a first value and a second value, the first value may indicate that the image is suitable for the task processing, and the second value may indicate that the image is not suitable for the task processing.

[0026] According to the seventh aspect, the circuitry can easily determine an image to be selected and an image not to be selected based on the first value and the second value.

[0027] In the decoder according to an eighth aspect of the present disclosure, in any one of the first to sixth aspects, the identifier may include a first value and a second value, the first value may indicate that the image is suitable for the task processing or that the image is not suitable for the task processing, and the second value may indicate that whether or not the image is suitable for the task processing is not specified.

[0028] According to the eighth aspect, the circuitry can easily determine an image to be selected

and an image not to be selected based on the first value. The ~~circuit~~ circuitry can arbitrarily determine, in accordance with its own processing capability or the like, whether or not to select an image associated with an identifier of the second value.

[0029] In the decoder according to a ninth aspect of the present disclosure, in any one of the first to eighth aspects, the image may include a plurality of images, the plurality of images may include an image not associated with the identifier, and the image not associated with the identifier may indicate an image that does not specify whether or not the image is suitable for the task processing.

[0030] According to the ninth aspect, it is possible to arbitrarily determine, in accordance with its own processing capability or the like, whether or not to select an image not associated with an identifier.

[0031] In the decoder according to a tenth aspect of the present disclosure, in any one of the first to ninth aspects, the identifier may include one or more identifiers, the task processing may include one or more types of task processing, and the circuitry may decode, from the bitstream, correspondence information indicating a correspondence between the one or more identifiers and the one or more types of task processing.

[0032] According to the tenth aspect, one or more pieces of task processing can be arbitrarily set in the correspondence information, and the execution accuracy of the task processing can be improved with respect to each of the one or more pieces of task processing having been set.

[0033] In the decoder according to an eleventh aspect of the present disclosure, in the tenth aspect, the correspondence information may be stored in a header area of the bitstream.

[0034] According to the eleventh aspect, the circuitry can easily acquire the correspondence information from the header area of the bitstream.

[0035] In the decoder according to a twelfth aspect of the present disclosure, in any one of the first to eleventh aspects, the identifier may be stored in a header area of the image associated with the identifier.

[0036] According to the twelfth aspect, the circuitry can easily acquire an identifier associated with each image from the header area of each image.

[0037] In the decoder according to a thirteenth aspect of the present disclosure, in any one of the first to eleventh aspect, the image may include a plurality of images, the identifier may include one or more identifiers, and the one or more identifiers may be stored in a header area of one image included in the plurality of images, and are associated with the one image and one or more images different from the one image.

[0038] According to the thirteenth aspect, the circuitry can easily acquire one or more identifiers associated with one or more images from the header area of one image.

[0039] In the decoder according to a fourteenth aspect of the present disclosure, in any one of the eleventh to thirteenth aspect, the header area may be an SEI area.

[0040] According to the fourteenth aspect, the circuitry can easily acquire an identifier from the SEI area.

[0041] In the decoder according to a fifteenth aspect of the present disclosure, in any one of the first to fourteenth aspect, the image may include a plurality of images, and the circuitry may decode, from the bitstream, specification information for specifying the image associated with the identifier among the plurality of images.

[0042] According to the fifteenth aspect, the number of identifiers and the number of images do not need to match by specifying, with the specification information, the image associated with the identifier, and therefore the data amount of the identifier can be reduced. This enables reduction in the code amount and improvement of the coding efficiency.

[0043] An encoder according to a sixteenth aspect of the present disclosure includes: ~~a circuit~~ circuitry; and a memory connected to the circuitry, in which the circuitry, in operation, encodes, to a bitstream, an image and an identifier associated with the image, the identifier indicating at least one of (i) the image being suitable for task processing corresponding to the identifier and (ii) the image being not suitable for the task processing.

[0044] According to the sixteenth aspect, the identifier indicates at least one of a fact that the image is suitable for task processing corresponding to the identifier and a fact that the image is not suitable for the task processing. Therefore, the decoder having received the bitstream can appropriately select, based on the identifier, at least one image to be used for the task processing from among the plurality of images decoded from the bitstream, whereby the execution accuracy of the task processing can be improved.

[0045] In the encoder according to a seventeenth aspect of the present disclosure, in the sixteenth aspect, the circuitry may set a value of the identifier based on an encoding condition under which the image is encoded.

[0046] According to the seventeenth aspect, after encoding an image, the circuitry can appropriately determine whether or not the image is suitable for task processing based on the encoding condition.

[0047] In the encoder according to an eighteenth aspect of the present disclosure, in the sixteenth or seventeenth aspect, the circuitry may set a value of the identifier based on an image quality of the image after encoding.

[0048] According to the eighteenth aspect, after encoding an image, the circuitry can appropriately determine whether or not the image is suitable for task processing based on the image quality.

[0049] In the encoding device according to a nineteenth aspect of the present disclosure, in any one of the sixteenth to eighteenth aspects, the circuitry may acquire, from an external device, an image of an encoding target and related information related to the image, and set a value of the identifier based on the related information.

[0050] According to the nineteenth aspect, the circuitry can appropriately determine whether or not the image is suitable for task processing based on the related information before encoding the image.

[0051] In the encoder according to a twentieth aspect of the present disclosure, in any one of the sixteenth to nineteenth aspect, the circuitry may perform image analysis on an image of an encoding target, and may set a value of the identifier based on a result of the image analysis.

[0052] According to the twentieth aspect, before encoding an image, the circuitry can appropriately determine whether or not the image is suitable for the task processing based on the result of the image analysis.

[0053] In the encoder according to a twenty-first aspect of the present disclosure, in any one of the sixteenth to twentieth aspects, the task processing may include a machine task.

[0054] According to the twenty-first aspect, the decoder having received the bitstream can select, based on the identifier, an image suitable for the machine task, whereby the machine task can be appropriately executed.

[0055] In the encoder according to a twenty-second aspect of the present disclosure, in any one of the sixteenth to twenty-first aspects, the task processing may include human vision.

[0056] According to the twenty-second aspect, the decoder having received the bitstream can select, based on the identifier, an image suitable for human vision, whereby human vision can be appropriately executed.

[0057] In the encoder according to a twenty-third aspect of the present disclosure, in any one of the sixteenth to twenty-second aspects, the identifier may include a plurality of identifiers, the task processing may include multiple types of task processing, and the plurality of identifiers may correspond to different types of task processing among the multiple types of task processing.

[0058] According to the twenty-third aspect, the plurality of identifiers can correspond to a plurality of pieces of task processing.

[0059] In the encoder according to a twenty-fourth aspect of the present disclosure, in the twenty-third aspect, the plurality of identifiers may include a first identifier and a second identifier, the task processing corresponding to the first identifier may include a machine task, and the task processing corresponding to the second identifier may include human vision.

[0060] According to the twenty-fourth aspect, the decoder having received the bitstream can select, based on the first identifier, an image suitable for the machine task, whereby the machine

task can be appropriately executed. The decoder can select, based on the second identifier, an image suitable for human vision, whereby human vision can be appropriately executed. That is, it becomes possible to appropriately execute the machine task or the human vision based on the first identifier and the second identifier according to the property of the image.

[0061] In the encoder according to a twenty-fifth aspect of the present disclosure, in any one of the sixteenth to twenty-fourth aspects, the identifier may include a first value and a second value, the first value may indicate that the image is suitable for the task processing, and the second value may indicate that the image is not suitable for the task processing.

[0062] According to the twenty-fifth aspect, the decoder having received the bitstream can easily determine an image to be selected and an image not to be selected based on the first value and the second value.

[0063] In the encoder according to a twenty-sixth aspect of the present disclosure, in any one of the sixteenth to twenty-fourth aspects, the identifier may include a first value and a second value, the first value may indicate that the image is suitable for the task processing or that the image is not suitable for the task processing, and the second value may indicate that whether or not the image is suitable for the task processing is not specified.

[0064] According to the twenty-sixth aspect, the decoder having received the bitstream can easily determine an image to be selected and an image not to be selected based on the first value. The decoder can arbitrarily determine, in accordance with its own processing capability or the like, whether or not to select an image associated with an identifier of the second value.

[0065] In the encoder according to a twenty-seventh aspect of the present disclosure, in any one of the sixteenth to twenty-sixth aspects, the image may include a plurality of images, the plurality of images may include an image not associated with the identifier, and the image not associated with the identifier may indicate an image that does not specify whether or not the image is suitable for the task processing.

[0066] According to the twenty-seventh aspect, the decoder having received the bitstream can arbitrarily determine, in accordance with its own processing capability or the like, whether or not to select an image not associated with an identifier.

[0067] In the encoder according to a twenty-eighth aspect of the present disclosure, in any one of the sixteenth to twenty-seventh aspect, the identifier may include one or more identifiers, the task processing may include one or more types of task processing, and the circuitry may further encode, to the bitstream, correspondence information indicating correspondence between the one or more identifiers and the one or more types of task processing.

[0068] According to the twenty-eighth aspect, one or more pieces of task processing can be arbitrarily set in the correspondence information, and the execution accuracy of the task processing

can be improved with respect to each of the one or more pieces of task processing having been set.

[0069] In the encoder according to a twenty-ninth aspect of the present disclosure, in any one of the sixteenth to twenty-eighth aspect, the correspondence information may be stored in a header area of the bitstream.

[0070] According to the twenty-ninth aspect, the decoder having received the bitstream can easily acquire the correspondence information from the header area of the bitstream.

[0071] In the encoder according to a thirtieth aspect of the present disclosure, in any one of the sixteenth to twenty-ninth aspects, the identifier may be stored in a header area of the image associated with the identifier.

[0072] According to the thirtieth aspect, the decoder having received the bitstream can easily acquire an identifier associated with each image from the header area of each image.

[0073] In the encoder according to a thirty-first aspect of the present disclosure, in any one of the sixteenth to thirtieth aspect, the image may include a plurality of images, the identifier may include one or more identifiers, and the one or more identifiers may be stored in a header area of one image included in the plurality of images, and are associated with the one image and one or more images different from the one image.

[0074] According to the thirty-first aspect, the decoder having received the bitstream can easily acquire one or more identifiers associated with one or more images from the header area of one image.

[0075] In the encoder according to a thirty-second aspect of the present disclosure, in any one of the twenty-ninth to thirty-first aspect, the header area may be an SEI area.

[0076] According to the thirty-second aspect, the circuitry can easily acquire an identifier from the SEI area.

[0077] In the encoder according to a thirty-third aspect of the present disclosure, in any one of the sixteenth to thirty-second aspect, the image may include a plurality of images, and the circuitry may further encode, to the bitstream, specification information for specifying the image associated with the identifier among the plurality of images.

[0078] According to the thirty-third aspect, the number of identifiers and the number of images do not need to match by specifying, with the specification information, the image associated with the identifier, and therefore the data amount of the identifier can be reduced. This enables reduction in the code amount and improvement of the coding efficiency.

[0079] A decoding method according to a thirty-fourth aspect of the present disclosure includes: by a decoder, decoding, from a bitstream, an image and an identifier associated with the image, the identifier indicating at least one of (i) the image being suitable for task processing corresponding to the identifier and (ii) the image being not suitable for the task processing.

[0080] According to the thirty-fourth aspect, the identifier indicates at least one of a fact that the image is suitable for task processing corresponding to the identifier and a fact that the image is not suitable for the task processing. Therefore, at least one image to be used for the task processing can be appropriately selected based on the identifier from among the plurality of images decoded from the bitstream, whereby the execution accuracy of the task processing can be improved.

[0081] An encoding method according to a thirty-fifth aspect of the present disclosure includes: by an encoder, encoding, to a bitstream, an image and an identifier associated with the image, the identifier indicating at least one of (i) the image being suitable for task processing corresponding to the identifier and (ii) the image being not suitable for the task processing.

[0082] According to the thirty-fifth aspect, the identifier indicates at least one of a fact that the image is suitable for task processing corresponding to the identifier and a fact that the image is not suitable for the task processing. Therefore, the decoder having received the bitstream can appropriately select, based on the identifier, at least one image to be used for the task processing from among the plurality of images decoded from the bitstream, whereby the execution accuracy of the task processing can be improved.

[0083] (Embodiments of Present Disclosure)

Embodiments of the present disclosure will be described in detail below with reference to the drawings. Note that elements denoted by the same reference numerals in different drawings represent the same or corresponding elements.

[0084] Note that each of the embodiments described below shows one specific example of the present disclosure. Numerical values, shapes, components, steps, order of steps, and the like shown in the following embodiments are merely examples, and are not intended to limit the present disclosure. A component that is not described in an independent claim representing the highest concept among components in the embodiments below is described as an arbitrary component. In all the embodiments, respective items of content can be combined.

[0085] Fig. 1 is a view illustrating a simplified configuration of an image processing system according to an embodiment of the present disclosure. The image processing system includes an encoder 1, a decoder 2, an image-capturing unit 3, a task processing unit 4, and a transmission path NW.

[0086] The image-capturing unit 3 is an example of an external device, and includes a camera that captures a moving image. The image-capturing unit 3 inputs image data D1 of a captured moving image to the encoder 1.

[0087] The encoder 1 transmits a bitstream BS to the decoder 2 via the transmission path NW. The decoder 2 receives the bitstream BS.

[0088] The task processing unit 4 executes task processing based on image data D2 input from the decoder 2. The task processing includes human vision and a machine task. Human vision is visual recognition or viewing of a moving image by a human such as an operator or a user. Machine tasks include various types of task processing such as object detection, object tracking, object segmentation, action recognition, or pose estimation using a machine-learned estimation model. The task processing unit 4 that executes human vision includes a display device such as a liquid crystal display or an organic EL display. The task processing unit 4 that executes a machine task includes an inferencer using AI.

[0089] The transmission path NW is the Internet, a wide area network (WAN), a local area network (LAN), or an arbitrary combination of them. The transmission path NW is desirably a private network or the like in which secure communication is secured by access restriction.

[0090] The encoder 1 includes circuitry 11 and a memory 12 connected to the circuitry 11. The circuitry 11 is configured to include a processor such as a CPU. The memory 12 is configured to include an arbitrary recording medium such as a ROM, a RAM, an HDD, an SSD, or a semiconductor memory. The memory 12 stores data or the like that is a processing target by the circuitry 11 or that is being processed.

[0091] The decoder 2 includes circuitry 21 and a memory 22 connected to the circuitry 21. The circuitry 21 is configured to include a processor such as a CPU. The memory 22 is configured to include an arbitrary recording medium such as a ROM, a RAM, an HDD, an SSD, or a semiconductor memory. The memory 22 stores data or the like that is a processing target by the circuitry 21 or that is being processed.

[0092] Fig. 2 is a view illustrating a simplified configuration of the circuitry 11 included in the encoder 1. The circuitry 11 includes a setting unit 31 and an encoding unit 32.

[0093] Fig. 3 is a flowchart showing the processing executed by the circuitry 11 included in the encoder 1.

[0094] First, in step SP11, the circuitry 11 acquires a plurality of images included in the image data D1 input from the image-capturing unit 3. The image data D1 is input to the setting unit 31 and the encoding unit 32.

[0095] Next, in step SP12, the setting unit 31 performs image analysis on each image of an encoding target included in the image data D1, and sets a value of an identifier based on a result of the image analysis. The setting unit 31 inputs, to the encoding unit 32, the identifier data D11 including the identifier whose value has been set.

[0096] The identifier indicates at least one of a fact that the image is suitable (i.e., appropriate) for the task processing corresponding to the identifier and a fact that the image is not suitable (i.e., inappropriate) for the task processing corresponding to the identifier. The fact that an image is

suitable for task processing includes a fact that use of the image is recommended when the task processing is executed on the decoder 2 side, or a fact that the image has been subjected to image processing suitable for the task processing. The fact that an image is not suitable for task processing includes a fact that use of the image is not recommended (i.e., unrecommended) when the task processing is executed on the decoder 2 side, or a fact that the image has not been subjected to image processing suitable for the task processing.

[0097] Next, in step SP13, the encoding unit 32 encodes, to the bitstream BS, the plurality of images included in the image data D1 input from the image-capturing unit 3 and a plurality of identifiers included in the identifier data D11 input from the setting unit 31.

[0098] Next, in step SP14, the circuitry 11 transmits the bitstream BS to the decoder 2 via the transmission path NW.

[0099] Fig. 4 is a view illustrating a first example of the image analysis by the setting unit 31. Fig. 4 illustrates three images in chronological order corresponding to times T_0, T_1, and T_2. The image corresponding to the time T_0 includes a tree and a pedestrian. The image corresponding to the time T_1 includes a tree and a pedestrian slightly moved from that at the time T_0. The image corresponding to the time T_2 includes a tree, a pedestrian slightly moving from that at the time T_1, and a jogger. In a case where the task processing corresponding to the identifier is human vision, the setting unit 31 sets all images to “appropriate” in order to perform smooth moving image reproduction.

[0100] Fig. 5 is a view illustrating a second example of the image analysis by the setting unit 31. In a case where the task processing corresponding to the identifier is object detection, since the pedestrian has already been detected at the time T_0, the setting unit 31 sets the images corresponding to the times T_0 and T_2 as “appropriate” and the image corresponding to the time T_1 as “inappropriate”.

[0101] Fig. 6 is a view illustrating a third example of the image analysis by the setting unit 31. In a case where the task processing corresponding to the identifier is object tracking, it is necessary to track the movement even if the movement amount is small, and therefore, the setting unit 31 sets all the images to “appropriate”.

[0102] The setting unit 31 may set the value of the identifier based on the image quality of the image before encoding. The setting unit 31 evaluates the image quality of the image before encoding on the basis of a sharpness value, a contrast value, a noise level, a peak signal to noise ratio (PSNR) value, or the like. The setting unit 31 may evaluate the image quality using a machine learning algorithm. By analyzing a histogram of the image, the setting unit 31 may evaluate the image quality depending on whether the image is overexposed, underexposed, or appropriately exposed.

[0103] In a case where the task processing is human vision, an image with low image quality deteriorates visibility. In a case where the task processing is object detection, an image with low image quality deteriorates estimation accuracy. Therefore, in a case where the task processing is human vision or object detection, the setting unit 31 may set an image whose image quality before encoding is equal to or greater than the threshold value as "appropriate", and may set an image whose image quality before encoding is less than the threshold value as "inappropriate".

[0104] In a case where the task processing is object tracking, the tolerance for a change in image quality is high, and therefore, the trajectory of an object can be detected even in an image with a low image quality. Therefore, in a case where the task processing is object tracking, the setting unit 31 may set all the images to "appropriate".

[0105] In this manner, the setting unit 31 determines whether the image is suitable (appropriate) or not suitable (inappropriate) for the task processing in accordance with the type of the task processing corresponding to the identifier. In the present embodiment, task processing executed by the task processing unit 4 is set to one specific type, and the encoder 1 and the decoder 2 share the setting information in advance.

[0106] Fig. 7 is a view illustrating the image data D1 and the identifier data D11. The image data D1 includes a plurality of images (five images P1 to P5 in this example) in chronological order. The identifier data D11 includes an identifier associated with each of the images P1 to P5. The identifier is 1-bit data, and the value of the identifier is set to "1" or "0".

[0107] Fig. 8 is a view illustrating a first example of syntax regarding setting of an identifier. The syntax includes an item of "frame_for_inference".

[0108] Fig. 9 is a view illustrating a first setting example of a value of an identifier in syntax. As illustrated in (A), in a case where the value of "frame_for_inference" is set to "1", it is indicated that the associated image is "appropriate". In a case where the value of "frame_for_inference" is set to "0", it is indicated that the associated image is "inappropriate".

[0109] Note that as illustrated in (B), in a case where the value of "frame_for_inference" is set to "0", it may be indicated that "appropriate" or "inappropriate" is not designated on the encoder 1 side, that is, whether or not the associated image is suitable for the task processing is not specified on the encoder 1 side.

[0110] Fig. 10 is a view illustrating a second example of syntax regarding setting of an identifier. The syntax includes an item of "frame_not_for_inference".

[0111] Fig. 11 is a view illustrating a second setting example of the value of the identifier in the syntax. As illustrated in (A), in a case where the value of "frame_not_for_inference" is set to "1", it is indicated that the associated image is "inappropriate". In a case where the value of "frame_not_for_inference" is set to "0", it is indicated that the associated image is "appropriate".

[0112] Note that as illustrated in (B), in a case where the value of “frame_not_for_inference” is set to “0”, it may be indicated that “appropriate” or “inappropriate” is not designated on the encoder 1 side, that is, whether or not the associated image is suitable for the task processing is not specified on the encoder 1 side.

[0113] Fig. 12 is a view illustrating a simplified configuration of packet data PD storing each of the encoded images P1 to P5. The packet data PD includes a header area 41 and a payload area 42. The header area 41 includes a supplemental enhancement information (SEI) area 43 for storing additional information. The encoding unit 32 stores encoding data of the image into the payload area 42, and stores the encoded data of the identifier associated with the image into the SEI area 43. However, the encoding unit 32 may store the encoded data of the identifier associated with the image into a VPS, an SPS, a PPS, a PH, an SH, an APS, a tile header, or the like, in place of the SEI area 43. The SEI area 43 may be prefix_SEI located before the payload area 42 or may be suffix_SEI located after the payload area 42.

[0114] Fig. 13 is a view illustrating a simplified configuration of the circuitry 21 included in the decoder 2. The circuitry 21 includes a decoding unit 51 and a selection unit 52.

[0115] Fig. 14 is a flowchart showing the processing executed by the circuitry 21 included in the decoder 2.

[0116] First, in step SP21, the circuitry 21 receives the bitstream BS from the encoder 1 via the transmission path NW. The bitstream BS is input to the decoding unit 51.

[0117] Next, in step SP22, by decoding the plurality of images and the identifier associated with each of the images from the bitstream BS, the circuitry 21 outputs image data D21 corresponding to the image data D1 and identifier data D22 corresponding to the identifier data D11. The image data D21 and the identifier data D22 are input to the selection unit 52.

[0118] Next, in step SP23, based on the identifier included in the identifier data D22, the selection unit 52 selects at least one image to be used for the task processing from among the plurality of images included in the image data D21.

[0119] Next, in step SP24, the selection unit 52 outputs the image data D2 including the at least one selected image. The image data D2 is input to the task processing unit 4. The task processing unit 4 executes the task processing using at least one image included in the input image data D2.

[0120] Fig. 15 is a view illustrating a first example of the image selection based on the identifier data D22. Fig. 15 illustrates an example in which, corresponding to (A) of Fig. 9, in a case where the value of “frame_for_inference” is set to “1”, it is indicated that the associated image is “appropriate”, and in a case where the value of “frame_for_inference” is set to “0”, it is indicated that the associated image is “inappropriate”.

[0121] The image data D21 includes a plurality of images (five images P1 to P5 in this example) in chronological order. The identifier data D22 includes an identifier associated with each of the images P1 to P5. The values of the identifiers are set to “1”, “0”, “0”, “1”, and “1” for the images P1 to P5, respectively.

[0122] The selection unit 52 selects the images P1, P4, and P5 whose values of the identifier are “1” among the plurality of images P1 to P5 included in the image data D21, and does not select the images P2 and P3 whose values of the identifier are “0”. As a result, the selection unit 52 outputs the image data D2 including the selected images P1, P4, and P5.

[0123] Fig. 16 is a view illustrating a second example of the image selection based on the identifier data D22. Fig. 16 illustrates an example in which, corresponding to (A) of Fig. 11, in a case where the value of “frame_not_for_inference” is set to “1”, it is indicated that the associated image is “inappropriate”, and in a case where the value of “frame_not_for_inference” is set to “0”, it is indicated that the associated image is “appropriate”.

[0124] The selection unit 52 does not select the images P1, P4, and P5 whose values of the identifier are “1” among the plurality of images P1 to P5 included in the image data D21, and selects the images P2 and P3 whose values of the identifier are “0”. As a result, the selection unit 52 outputs the image data D2 including the selected images P2 and P3.

[0125] Note that the designation of the image by the identifier is merely a proposal on whether the selection is appropriate or inappropriate, and does not force selection or non-selection by the selection unit 52. Therefore, depending on the processing capability of the hardware resource of the decoder 2 or the like, the selection unit 52 needs not select an image determined as appropriate in a case where there is no available capacity, or may select an image determined as inappropriate in a case where there is an available capacity.

[0126] According to the encoder 1 according to the present embodiment, the identifier indicates at least one of a fact that the image is suitable for task processing corresponding to the identifier and a fact that the image is not suitable for the task processing. Therefore, the decoder 2 having received the bitstream BS can appropriately select, based on the identifier, at least one image to be used for the task processing from among the plurality of images decoded from the bitstream BS, whereby the execution accuracy of the task processing can be improved.

[0127] According to the encoder 1 according to the present embodiment, the ~~circuit~~ circuitry 11 can appropriately determine whether or not the image is suitable for the task processing based on a result of image analysis before encoding the image.

[0128] According to the encoder 1 according to the present embodiment, the task processing includes a machine task. Therefore, the decoder 2 having received the bitstream BS can select, based on the identifier, an image suitable for the machine task, whereby the machine task can be

appropriately executed.

[0129] According to the encoder 1 according to the present embodiment, the task processing includes human vision. Therefore, the decoder 2 having received the bitstream BS can select, based on the identifier, an image suitable for human vision, whereby human vision can be appropriately executed.

[0130] According to the encoder 1 according to the present embodiment, the identifier includes the first value and the second value, where the first value indicates that the image is suitable for the task processing, and the second value indicates that the image is not suitable for the task processing. Therefore, the decoder 2 having received the bitstream BS can easily determine an image to be selected and an image not to be selected based on the first value and the second value.

[0131] According to the encoder 1 according to the present embodiment, the identifier includes the first value and the second value, where the first value indicates that the image is suitable for the task processing or the image is not suitable for the task processing, and the second value indicates that whether or not the image is suitable for the task processing is not specified. Therefore, the decoder 2 having received the bitstream BS can easily determine an image to be selected and an image not to be selected based on the first value. The decoder 2 can arbitrarily determine, in accordance with its own processing capability or the like, whether or not to select an image associated with an identifier of the second value.

[0132] According to the encoder 1 according to the present embodiment, the identifier is stored in the header area 41 of the image associated with the identifier. Therefore, the decoder 2 having received the bitstream BS can easily acquire an identifier associated with each image from the header area 41 of each image.

[0133] According to the decoder 2 according to the present embodiment, the identifier indicates at least one of a fact that the image is suitable for task processing corresponding to the identifier and a fact that the image is not suitable for the task processing. Therefore, at least one image to be used for the task processing can be appropriately selected based on the identifier from among the plurality of images decoded from the bitstream BS, whereby the execution accuracy of the task processing can be improved.

[0134] According to the decoder 2 according to the present embodiment, an image suitable for the task processing can be selected based on the identifier, whereby the execution accuracy of the task processing can be improved.

[0135] According to the decoder 2 according to the present embodiment, the task processing includes a machine task. Therefore, an image suitable for the machine task can be selected based on the identifier, whereby the machine task can be appropriately executed.

[0136] According to the decoder 2 according to the present embodiment, the task processing

includes human vision. Therefore, an image suitable for human vision can be selected based on the identifier, whereby human vision can be appropriately executed.

[0137] According to the decoder 2 according to the present embodiment, the identifier includes the first value and the second value, where the first value indicates that the image is suitable for the task processing, and the second value indicates that the image is not suitable for the task processing. Therefore, the circuitry 21 can easily determine an image to be selected and an image not to be selected based on the first value and the second value.

[0138] According to the decoder 2 according to the present embodiment, the identifier includes the first value and the second value, where the first value indicates that the image is suitable for the task processing or the image is not suitable for the task processing, and the second value indicates that whether or not the image is suitable for the task processing is not specified. Therefore, the circuitry 21 can easily determine an image to be selected and an image not to be selected based on the first value. The circuitry 21 can arbitrarily determine, in accordance with its own processing capability or the like, whether or not to select an image associated with an identifier of the second value.

[0139] According to the decoder 2 according to the present embodiment, the identifier is stored in the header area 41 of the image associated with the identifier. Therefore, the circuitry 21 can easily acquire an identifier associated with each image from the header area 41 of each image.

[0140] Various modifications of the embodiment of the present disclosure will be described below. The modifications can be applied in any combination.

[0141] (First Modification)

In the above embodiment, each identifier is stored in the header area 41 of an image associated with the identifier, but the present invention is not limited to this. The circuitry 11 may define an image group including a plurality of images, and collectively store, into the header area 41 of one image (e.g., head image) in the image group, a plurality of identifiers associated with the plurality of images included in the image group. The image group may be a group of pictures (GOP) including an I picture, a P picture, and a B picture, or may be a set of a plurality of arbitrary images.

[0142] Fig. 17 is a view illustrating the image data D1 and the identifier data D11. The image data D1 includes a plurality of images (32 images P1 to P32 in this example) in chronological order. The images P1 to P32 are included in the same image group. The identifier data D11 includes 32 identifiers associated with the 32 images P1 to P32. An identifier set including the 32 identifiers is stored in the header area 41 of the image P1 that is a head of the image group.

[0143] Although one identifier and one image are associated with each other in Fig. 17, one identifier and a plurality of images may be associated with each other. For example, when a

plurality of identifiers having the same value are consecutive, one identifier of the head of the plurality of identifiers may be associated with a plurality of images corresponding to the plurality of identifiers.

[0144] Fig. 18 is a view illustrating an example of syntax regarding setting of an identifier. The syntax includes an item of "num_frames_minus1". The value of "num_frames_minus1" is set to a value in which "1" is subtracted from the number of images included in the image group. In a case where the image group includes the 32 images P1 to P32, the value of "num_frames_minus1" is set to "31".

[0145] The value of "frame_for_inference" related to each of the plurality of images included in the image group is set by loop processing indicated by a for statement. At this time, which image each of the plurality of "frame_for_inference" is associated with may be determined according to the order of the loop processing indicated by the for statement.

[0146] In place of the description using a repeated statement such as a for statement, a plurality of setting values may be described in itemized form. The same applies to the following modifications.

[0147] According to the present modification, the decoder 2 having received the bitstream BS can easily acquire the plurality of identifiers associated with the plurality of images included in the image group from the header area 41 of one image.

[0148] (Second Modification)

In a first modification, the same number of identifiers as the plurality of images included in the image group are collectively stored in the header area 41 of one image in the image group, but the present invention is not limited to this. The circuitry 11 may further encode, to the bitstream BS, specification information for specifying the image associated with the identifier among the plurality of images included in the image group.

[0149] Fig. 19 is a view illustrating an example of syntax regarding setting of an identifier. The syntax includes an item of "image_identifier[i]" indicating specification information. The value of "image_identifier[i]" is set by using a picture order count (POC), a difference value of the POC, the number of counts, an ID of a sequence parameter set (SPS), an ID of a picture parameter set (PPS), or the like.

[0150] The value of "frame_for_inference" is set for the image specified by the specification information among the plurality of images included in the image group.

[0151] Note that an image not specified by the specification information, that is, an image not associate with the identifier indicates that whether or not the image is suitable for the task processing is not specified.

[0152] In a case where the identifier set is not included in the image of the head in the image

group, it is indicated that whether or not the image is suitable for the task processing is not specified for all the images included in the group.

[0153] According to the present modification, by specifying, with the specification information, the image associated with the identifier, it is not necessary to match the number of identifiers included in the identifier set with the number of images included in the image group, and therefore the data amount of the identifier can be reduced.

[0154] (Third Modification)

In the above embodiment, the identifiers are associated with all the images, but the present invention is not limited to this. The circuitry 11 needs not associate an identifier with an image that does not specify whether or not the image is suitable for task processing.

[0155] Fig. 20 is a view illustrating a first setting example of the value of the identifier in the syntax, corresponding to Fig. 9. As illustrated in (A), in a case where the value of "frame_for_inference" is set to "1", it is indicated that the associated image is "appropriate". In a case where the value of "frame_for_inference" is set to "0", it is indicated that the associated image is "inappropriate". The image not associated with an identifier indicates that whether or not the image is suitable for task processing is not specified on the encoder 1 side.

[0156] Note that as illustrated in (B), in a case where the value of "frame_for_inference" is set to "0", or regarding an image not associated with an identifier, it may be indicated that whether or not the image is suitable for the task processing is not specified on the encoder 1 side.

[0157] Fig. 21 is a view illustrating a second setting example of the value of the identifier in the syntax, corresponding to Fig. 10. As illustrated in (A), in a case where the value of "frame_not_for_inference" is set to "1", it is indicated that the associated image is "inappropriate". In a case where the value of "frame_not_for_inference" is set to "0", it is indicated that the associated image is "appropriate". The image not associated with an identifier indicates that whether or not the image is suitable for task processing is not specified on the encoder 1 side.

[0158] Note that as illustrated in (B), in a case where the value of "frame_not_for_inference" is set to "0", or regarding an image not associated with an identifier, it may be indicated that whether or not the image is suitable for the task processing is not specified on the encoder 1 side.

[0159] Fig. 22 is a view illustrating an example of syntax regarding setting of an identifier on the basis of Fig. 18. Fig. 23 is a view illustrating an example of syntax regarding setting of an identifier on the basis of Fig. 19. The syntax includes an item of "frame_for_inference_present_flag[i]". For an image associated with an identifier, the value of "frame_for_inference_present_flag[i]" is set to "1". For an image not associated with an identifier, the value of "frame_for_inference_present_flag[i]" is set to "0".

[0160] (Fourth Modification)

In the above embodiment, the task processing executed by the task processing unit 4 is set to one specific type, but the present invention is not limited to this. The circuitry 11 may further encode, to the bitstream BS, setting information for arbitrarily setting the type of task processing. The setting information may be stored in the SEI area included in the header area of the bitstream BS.

[0161] Fig. 24 is a view illustrating an example of syntax regarding setting of task processing. The syntax includes an item of "vision_task". In the item of "vision_task", a character string indicating the type of task processing is described. Alternatively, in place of the character string, an integer value indicating the type of task processing may be described. In this case, the correspondence relationship between the integer value and the type of task processing is set, and the encoder 1 and the decoder 2 share the setting information in advance.

[0162] According to the present modification, the type of task processing corresponding to the identifier can be arbitrarily set.

[0163] (Fifth Modification)

In the above embodiment, the setting unit 31 sets the value of the identifier based on the result of the image analysis on the image before encoding, but the present invention is not limited to this. The setting unit 31 may set the value of the identifier based on the image after encoding.

[0164] Fig. 25 is a view illustrating a simplified configuration of the circuitry 11 included in the encoder 1. The circuitry 11 includes a setting unit 31 and an encoding unit 32. The image data D1 including a plurality of images before encoding and image data D12 including a plurality of images encoded by the encoding unit 32 are input to the setting unit 31. Based on the image data D1 and D12, the setting unit 31 derives an encoding condition when the encoding unit 32 encodes each image. The encoding condition includes conditions related to the image quality of the image after encoding, such as a value of a quantization parameter, a PSNR value of the image after encoding, or a hierarchical level of a temporal layer. The larger the value of the quantization parameter is, the lower the image quality of the image after encoding is. The smaller the PSNR value is, the lower the image quality of the image after encoding is. The higher the hierarchical level of the temporal layer is, the lower the image quality of the image after encoding is.

[0165] In a case where the task processing is human vision, an image with low image quality deteriorates visibility. In a case where the task processing is object detection, an image with low image quality deteriorates estimation accuracy. Therefore, in a case where the task processing is human vision or object detection, the setting unit 31 may set an image whose image quality after encoding is equal to or greater than the threshold value as "appropriate", and may set an image whose image quality after encoding is less than the threshold value as "inappropriate".

[0166] In a case where the task processing is object tracking, the tolerance for a change in image

quality is high, and therefore, the trajectory of an object can be detected even in an image with a low image quality. Therefore, in a case where the task processing is object tracking, the setting unit 31 may set all the images to “appropriate”.

[0167] The setting unit 31 inputs, to the encoding unit 32, the identifier data D11 including an identifier whose value is set based on the image after encoding.

[0168] According to the present modification, after encoding an image, the circuitry 11 can appropriately determine whether or not the image is suitable for task processing based on the encoding condition.

[0169] According to the present modification, after encoding an image, the ~~circuit~~ circuitry 11 can appropriately determine whether or not the image is suitable for task processing based on the image quality.

[0170] (Sixth Modification)

In the above embodiment, the setting unit 31 sets the value of the identifier based on the result of the image analysis on the image before encoding, but the present invention is not limited to this. The setting unit 31 may set the value of the identifier based on related information related to the image before encoding.

[0171] Fig. 26 is a view illustrating a simplified configuration of the circuitry 11 included in the encoder 1. The circuitry 11 includes a setting unit 31 and an encoding unit 32. The image data D1 including a plurality of images before encoding and related information D13 detected by the image-capturing unit 3 when the image-capturing unit 3 captures a moving image are input to the setting unit 31 from the image-capturing unit 3.

[0172] The camera included in the image-capturing unit 3 has various built-in sensors such as a proximity sensor, a gyro sensor, an acceleration sensor, and a geomagnetic sensor. The related information D13 includes detection values detected by these sensors.

[0173] The setting unit 31 sets a value of the identifier based on the image data D1 and the related information D13.

[0174] In a case where an object located immediately in front of the camera is detected by the proximity sensor and this state continues for a certain period of time or more, the setting unit 31 may set an image captured during this period as “inappropriate”. In a case where an object located immediately in front of the camera is detected by the proximity sensor and the object is moving at a high speed of a certain speed or faster, the setting unit 31 may set an image captured during this period as “appropriate”.

[0175] In a case where camera shake is detected by a gyro sensor, an acceleration sensor, a geomagnetic sensor, or the like, and the level of the shake is equal to or greater than a threshold, the setting unit 31 may set an image captured during that time as “inappropriate”.

[0176] According to the present modification, the circuitry 11 can appropriately determine whether or not the image is suitable for task processing based on the related information D13 before encoding the image.

[0177] (Seventh Modification)

In the above embodiment, there is one type of task processing, but the present invention is not limited to this. The task processing may include a plurality of types of task processing, and the circuitry 11 may individually set "appropriate" or "inappropriate" of the image for each task processing. Hereinafter, an explanation will be given regarding an example in which the task processing includes first task processing and second task processing, the first task processing is object detection, and the second task processing is human vision. However, the task processing may include three or more types of task processing, and the type of task processing is not limited to object detection and human vision.

[0178] Fig. 27 is a view illustrating a simplified configuration of the image processing system according to the seventh modification. The image processing system includes a first task processing unit 4A and a second task processing unit 4B. The first task processing unit 4A executes the first task processing that is object detection based on first image data D2A input from the decoder 2. The second task processing unit 4B executes the second task processing that is human vision based on second image data D2B input from the decoder 2.

[0179] Fig. 28 is a view illustrating an example of setting information 60 of task processing. The setting information 60 indicates that the first task processing is object detection and the second task processing is human vision. The encoder 1 and the decoder 2 share the setting information 60 in advance.

[0180] Fig. 29 is a view illustrating a simplified configuration of the circuitry 11 included in the encoder 1. The circuitry 11 includes a setting unit 31 and an encoding unit 32. The setting unit 31 inputs, to the encoding unit 32, first identifier data D11A including a plurality of first identifiers related to the first task processing, and second identifier data D11B including a plurality of second identifiers related to the second task processing.

[0181] Fig. 30 is a flowchart showing the processing executed by the circuitry 11 included in the encoder 1.

[0182] First, in step SP11, the circuitry 11 acquires a plurality of images included in the image data D1 input from the image-capturing unit 3. The image data D1 is input to the setting unit 31 and the encoding unit 32.

[0183] Next, in step SP12A, the setting unit 31 performs image analysis on each image of an encoding target included in the image data D1, and sets a value of the first identifier and a value of the second identifier based on a result of the image analysis. The setting unit 31 inputs, to the

encoding unit 32, the first identifier data D11A including the first identifier whose value has been set and the second identifier data D11B including the second identifier whose value has been set.

[0184] Next, in step SP13A, the encoding unit 32 encodes, to the bitstream BS, the plurality of images included in the image data D1 having been input from the image-capturing unit 3, the plurality of first identifiers included in the first identifier data D11A having been input from the setting unit 31, and the plurality of second identifiers included in the second identifier data D11B having been input from the setting unit 31.

[0185] Next, in step SP14, the circuitry 11 transmits the bitstream BS to the decoder 2 via the transmission path NW.

[0186] Fig. 31 is a view illustrating the image data D1, first identifier data D11A, and second identifier data D11B. The image data D1 includes a plurality of images (five images P1 to P5 in this example) in chronological order. The first identifier data D11A and the second identifier data D11B include the first identifier and the second identifier associated with each of the images P1 to P5. Each of the first identifier and the second identifier is 1-bit data, and each value of the first identifier and the second identifier is set to “1” or “0”.

[0187] Fig. 32 is a view illustrating a bit string including the first identifier data D11A and the second identifier data D11B. The two bits of the head of the bit string indicate the respective values of the first identifier and the second identifier associated with the image P1. The subsequent two bits of the bit string indicate the respective values of the first identifier and the second identifier associated with the image P2. Similarly, consecutive two bits of the bit string indicate the respective values of the first identifier and the second identifier associated with the images P3 to P5. Note that in a case where the number of types of the task processing is N, consecutive N bits of the bit string indicate the respective values of the N identifiers associated with the same image.

[0188] Fig. 33A is a view illustrating a first example of syntax regarding setting of the first identifier and the second identifier. The syntax includes an item of “num_vision_task_minus1”. The value of “num_vision_task_minus1” is set to a value in which “1” is subtracted from the number of types of the task processing. The value of “frame_for_inference[i]” when i = 1 corresponds to a value of the first identifier, and the value of “frame_for_inference[i]” when i = 2 corresponds to a value of the second identifier. In a case where the value of “frame_for_inference[i]” is set to “1”, it is indicated that the associated image is “appropriate”. In a case where the value of “frame_for_inference[i]” is set to “0”, it is indicated that the associated image is “inappropriate”.

[0189] In place of the description using a repeated statement such as a for statement, a plurality of setting values of “frame_for_inference[i]” may be described in itemized form as illustrated in

Fig. 33B.

[0190] Fig. 34A is a view illustrating a second example of syntax regarding setting of the first identifier and the second identifier. The value of "frame_not_for_inference[i]" when i = 1 corresponds to a value of the first identifier, and the value of "frame_not_for_inference[i]" when i = 2 corresponds to a value of the second identifier. In a case where the value of "frame_not_for_inference[i]" is set to "1", it is indicated that the associated image is "inappropriate". In a case where the value of "frame_not_for_inference[i]" is set to "0", it is indicated that the associated image is "appropriate".

[0191] In place of the description using a repeated statement such as a for statement, a plurality of setting values of "frame_not_for_inference[i]" may be described in itemized form as illustrated in Fig. 34B.

[0192] Fig. 35 is a view illustrating a simplified configuration of the circuitry 21 included in the decoder 2. The circuitry 21 includes a decoding unit 51 and a selection unit 52. The decoding unit 51 inputs, to the selection unit 52, the image data D21, first identifier data D22A including a plurality of first identifiers related to the first task processing, and second identifier data D22B including a plurality of second identifiers related to the second task processing.

[0193] Fig. 36 is a flowchart showing the processing executed by the ~~circuit~~ circuitry 21 included in the decoder 2.

[0194] First, in step SP21, the circuitry 21 receives the bitstream BS from the encoder 1 via the transmission path NW. The bitstream BS is input to the decoding unit 51.

[0195] Next, in step SP22A, by decoding, from the bitstream BS, the plurality of images and the first identifier and the second identifier associated with each image, the circuitry 21 outputs the image data D21 corresponding to the image data D1, the first identifier data D22A corresponding to the first identifier data D11A, and the second identifier data D22B corresponding to the second identifier data D11B. The image data D21, the first identifier data D22A, and the second identifier data D22B are input to the selection unit 52.

[0196] Next, in step SP23A, based on the first identifier included in the first identifier data D22A, the selection unit 52 selects at least one first image to be used for the first task processing from among the plurality of images included in the image data D21. Based on the second identifier included in the second identifier data D22B, the selection unit 52 selects at least one second image to be used for the second task processing from among the plurality of images included in the image data D21.

[0197] Next, in step SP24A, the selection unit 52 outputs the first image data D2A including the at least one first image having been selected and the second image data D2B including the at least one second image having been selected. The first image data D2A is input to the first task

processing unit 4A, and the second image data D2B is input to the second task processing unit 4B. The first task processing unit 4A executes the first task processing by using at least one first image included in the input first image data D2A. The second task processing unit 4B executes the second task processing by using at least one second image included in the input second image data D2B.

[0198] Fig. 37 is a view illustrating an example of selection of the first image and the second image based on the first identifier data D22A and the second identifier data D22B. Fig. 37 illustrates an example in which, in a case where the value of "frame_for_inference[i]" is set to "1", it is indicated that the associated image is "appropriate", and in a case where the value of "frame_for_inference[i]" is set to "0", it is indicated that the associated image is "inappropriate".

[0199] The image data D21 includes a plurality of images (five images P1 to P5 in this example) in chronological order. The first identifier data D22A includes the first identifier associated with each of the images P1 to P5. The values of the first identifiers are set to "1", "0", "0", "1", and "1" for the images P1 to P5, respectively. The second identifier data D22B includes the second identifier associated with each of the images P1 to P5. The values of the second identifiers are set to "1" for all the images P1 to P5.

[0200] The selection unit 52 selects the images P1, P4, and P5 whose values of the first identifier are "1" among the plurality of images P1 to P5 included in the image data D21, and does not select the images P2 and P3 whose values of the first identifier are "0". As a result, the selection unit 52 outputs the first image data D2A including the selected images P1, P4, and P5.

[0201] The selection unit 52 selects the images P1 to P5 whose values of the second identifier are "1" among the plurality of images P1 to P5 included in the image data D21. As a result, the selection unit 52 outputs the second image data D2B including all the images P1 to P5.

[0202] According to the present modification, the first image suitable for the machine task can be selected based on the first identifier, whereby the machine task can be appropriately executed. The second image suitable for human vision can be selected based on the second identifier, whereby human vision can be appropriately executed.

[0203] (Eighth Modification)

In the above embodiment, each identifier is 1-bit data, but the present invention is not limited to this. Each identifier may be multiple-bit data.

[0204] Fig. 38 is a view illustrating a setting example of a value of an identifier in syntax. In this example, each identifier is 2-bit data.

[0205] In a case where the value of "frame_for_inference" is set to "0", it is indicated that "appropriate" or "inappropriate" is not designated on the encoder 1 side, that is, whether or not the associated image is suitable for task processing is not specified on the encoder 1 side. In a case

where the value of “frame_for_inference” is set to “1”, it is indicated that the associated image is “inappropriate”. In a case where the value of “frame_for_inference” is set to “2”, it is indicated that the associated image is “appropriate”. Note that the value of “3” in “frame_for_inference” is assigned for future expansion and is currently undefined.

[0206] According to the present modification, since three states of “appropriate”, “inappropriate”, and “not designated” can be designated by each 2-bit identifier, it is not necessary to mix an image associated with an identifier and an image not associated with an identifier.

[0207] (Ninth Modification)

In a ninth modification, a combination of the third modification and the seventh modification will be described.

[0208] Fig. 39 is a view illustrating an example of selection of the first image and the second image based on the first identifier data D22A and the second identifier data D22B. Fig. 39 illustrates an example in which, in a case where the value of “frame_for_inference[i]” is set to “1”, it is indicated that the associated image is “appropriate”, in a case where the value of “frame_for_inference[i]” is set to “0”, it is indicated that the associated image is “inappropriate”, and regarding an image not associated with the first identifier and the second identifier, it is indicated that whether or not the image is suitable for the task processing is not specified on the encoder 1 side.

[0209] The selection unit 52 selects, among the plurality of images P1 to P32 included in the image data D21, the images P2 and P32 whose values of the first identifier are “1” and the image P1 not associated with the first identifier and the second identifier. The selection unit 52 does not select the image P4 whose values of the first identifier are “0” and the image P3 not associated with the first identifier and the second identifier. As a result, the selection unit 52 outputs the first image data D2A including the selected images P1, P2, and P32.

[0210] The selection unit 52 selects, among the plurality of images P1 to P32 included in the image data D21, the images P4 and P32 whose values of the second identifier are “1” and the image P1 not associated with the first identifier and the second identifier. The selection unit 52 does not select the image P2 whose values of the second identifier are “0” and the image P3 not associated with the first identifier and the second identifier. As a result, the selection unit 52 outputs the second image data D2B including the selected images P1, P4, and P32.

[0211] (Tenth Modification)

In the seventh modification, the first task processing executed by the first task processing unit 4A and the second task processing executed by the second task processing unit 4B are set in advance by the setting information 60, but the present invention is not limited to this. The ~~circuit~~ circuitry 11 may further encode, to the bitstream BS, correspondence information indicating

correspondence between the first identifier and the second identifier and the first task processing and the second task processing.

[0212] Fig. 40 is a view illustrating an example of syntax regarding setting of task processing. The syntax includes an item of "num_identifier_minus1". The value of "num_identifier_minus1" is set to a value in which "1" is subtracted from the number of identifiers associated with the image. The syntax includes an item of "vision_task[i]". In the item of "vision_task[i]", a character string indicating the type of task processing is described. Alternatively, in place of the character string, an integer value indicating the type of task processing may be described. In this case, the correspondence relationship between the integer value and the type of task processing is set, and the encoder 1 and the decoder 2 share the setting information in advance.

[0213] Fig. 41 is a view illustrating a simplified configuration of the bitstream BS. The bitstream BS includes a header area 71 and a payload area 72. The header area 71 includes an SEI area 73 for storing additional information. The encoding unit 32 stores, into the SEI area 73, encoding data of the correspondence information indicating the correspondence between the plurality of identifiers and the multiple types of task processing. However, the encoding unit 32 may store the encoded data of the identifier associated with the image into a VPS, an SPS, a PPS, a PH, an SH, an APS, a tile header, or the like, in place of the SEI area 73. The SEI area 73 may be prefix_SEI located before the payload area 72 or may be suffix_SEI located after the payload area 72.

[0214] According to the present modification, multiple types of task processing can be arbitrarily set in the correspondence information, and the execution accuracy of the task processing can be improved with respect to each of the multiple types of task processing having been set.

[0215] (Eleventh Modification)

In an eleventh modification, a combination of the first modification and the seventh modification will be described.

[0216] Fig. 42 is a view illustrating an example of syntax regarding setting of the first identifier and the second identifier. In a case where the value of "frame_for_inference[i][j]" is set to "1", it is indicated that the i-th image in the image group is "appropriate" for the j-th task processing. In a case where the value of "frame_for_inference[i][j]" is set to "0", it is indicated that the i-th image in the image group is "inappropriate" for the j-th task processing.

[0217] (Twelfth Modification)

In a twelfth modification, a combination of the second modification and the seventh modification will be described.

[0218] Fig. 43 is a view illustrating an example of syntax regarding setting of the first identifier and the second identifier. In a case where the value of "frame_for_inference[i][j]" is set to "1",

it is indicated that the i-th image specified by the specification information in the image group is "appropriate" for the j-th task processing. In a case where the value of "frame_for_inference[i][j]" is set to "0", it is indicated that the i-th image specified by the specification information in the image group is "inappropriate" for the j-th task processing.

[0219] (Thirteenth Modification)

In a thirteenth modification, a combination of the third modification and the seventh modification will be described.

[0220] Fig. 44 is a view illustrating an example of syntax regarding setting of the first identifier and the second identifier. For an image associated with an identifier, the value of "frame_for_inference_present_flag[i]" is set to "1". For an image not associated with an identifier, the value of "frame_for_inference_present_flag[i]" is set to "0".

[0221] Regarding an image associated with an identifier, in a case where the value of "frame_for_inference[i][j]" is set to "1", it is indicated that the i-th image in the image group is "appropriate" for the j-th task processing. In a case where the value of "frame_for_inference[i][j]" is set to "0", it is indicated that the i-th image in the image group is "inappropriate" for the j-th task processing.

[0222] (Fourteenth Modification)

In a fourteenth modification, a combination of the second modification, the third modification, and the seventh modification will be described.

[0223] Fig. 45 is a view illustrating an example of syntax regarding setting of the first identifier and the second identifier. Regarding an image associated with an identifier, in a case where the value of "frame_for_inference[i][j]" is set to "1", it is indicated that the i-th image specified by the specification information in the image group is "appropriate" for the j-th task processing. In a case where the value of "frame_for_inference[i][j]" is set to "0", it is indicated that the i-th image specified by the specification information in the image group is "inappropriate" for the j-th task processing.

[0224] The present disclosure is particularly useful for application to an image processing system including an encoder that encodes an image to a bitstream and transmits the bitstream, and a decoder that decodes an image from a received bitstream.

CLAIMS

1. A decoder comprising:

circuitry; and

a memory connected to the circuitry,

wherein the circuitry, in operation:

decodes, from a bitstream, an image and an identifier associated with the image, the identifier indicating at least one of (i) the image being suitable for task processing corresponding to the identifier and (ii) the image being not suitable for the task processing.

2. The decoder according to claim 1, wherein

the image includes a plurality of images, and

the circuitry selects, based on the identifier, at least one image to be used for the task processing from the plurality of images.

3. The decoder according to claim 1, wherein the task processing includes a machine task.

4. The decoder according to claim 1, wherein the task processing includes human vision.

5. The decoder according to claim 1, wherein

the identifier includes a plurality of identifiers,

the task processing includes multiple types of task processing, and

the plurality of identifiers correspond to different types of task processing among the multiple types of task processing.

6. The decoder according to claim 5, wherein

the plurality of identifiers include a first identifier and a second identifier,

the task processing corresponding to the first identifier includes a machine task, and

the task processing corresponding to the second identifier includes human vision.

7. The decoder according to claim 1, wherein

the identifier includes a first value and a second value, the first value indicating that the image is suitable for the task processing, and the second value indicating that the image is not suitable for the task processing.

8. The decoder according to claim 1, wherein

the identifier includes a first value and a second value, the first value indicating that either the image is suitable for the task processing or that the image is not suitable for the task processing, and the second value indicating that whether or not the image is suitable for the task processing is not specified.

9. The decoder according to claim 1, wherein

the image includes a plurality of images, the plurality of images including an image not associated with the identifier, and the image not associated with the identifier indicating an image that does not specify whether or not the image is suitable for the task processing.

10. The decoder according to claim 1, wherein

the identifier includes one or more identifiers,

the task processing includes one or more types of task processing, and

the circuitry decodes, from the bitstream, correspondence information indicating a correspondence between the one or more identifiers and the one or more types of task processing.

11. The decoder according to claim 10, wherein the correspondence information is stored in a header area of the bitstream.

12. The decoder according to claim 1, wherein the identifier is stored in a header area of the image associated with the identifier.

13. The decoder according to claim 1, wherein

the image includes a plurality of images,

the identifier includes one or more identifiers, and

the one or more identifiers are stored in a header area of one image included in the plurality of images, and are associated with the one image and one or more images different from the one image.

14. The decoder according to any one of claims 11 to 13, wherein the header area is an SEI area.

15. The decoder according to claim 1, wherein

the image includes a plurality of images, and

the circuitry decodes, from the bitstream, specification information for specifying the image associated with the identifier among the plurality of images.

16. An encoder comprising:
circuitry; and
a memory connected to the circuitry,
wherein the circuitry, in operation:
encodes, to a bitstream, an image and an identifier associated with the image, the identifier indicating at least one of (i) the image being suitable for task processing corresponding to the identifier and (ii) the image being not suitable for the task processing.

17. The encoder according to claim 16, wherein the circuitry sets a value of the identifier based on an encoding condition under which the image is encoded.

18. The encoder according to claim 16, wherein the circuitry sets a value of the identifier based on an image quality of the image after encoding.

19. The encoder according to claim 16, wherein
the circuitry
acquires, from an external device, an image of an encoding target and related information related to the image, and
sets a value of the identifier based on the related information.

20. The encoder according to claim 16, wherein
the circuitry
performs image analysis on an image of an encoding target, and
sets a value of the identifier based on a result of the image analysis.

21. The encoder according to claim 16, wherein the task processing includes a machine task.

22. The encoder according to claim 16, wherein the task processing includes human vision.

23. The encoder according to claim 16, wherein
the identifier includes a plurality of identifiers,
the task processing includes multiple types of task processing, and

the plurality of identifiers correspond to different types of task processing among the multiple types of task processing.

24. The encoder according to claim 23, wherein

the plurality of identifiers include a first identifier and a second identifier,

the task processing corresponding to the first identifier includes a machine task, and

the task processing corresponding to the second identifier includes human vision.

25. The encoder according to claim 16, wherein

the identifier includes a first value and a second value, the first value indicating that the image is suitable for the task processing, and the second value indicating that the image is not suitable for the task processing.

26. The encoder according to claim 16, wherein

the identifier includes a first value and a second value, the first value indicating that the image is suitable for the task processing or that the image is not suitable for the task processing, and the second value indicating that whether or not the image is suitable for the task processing is not specified.

27. The encoder according to claim 16, wherein

the image includes a plurality of images, the plurality of images including an image not associated with the identifier, and the image not associated with the identifier indicating an image that does not specify whether or not the image is suitable for the task processing.

28. The encoder according to claim 16, wherein

the identifier includes one or more identifiers,

the task processing includes one or more types of task processing, and

the circuitry encodes, to the bitstream, correspondence information indicating correspondence between the one or more identifiers and the one or more types of task processing.

29. The encoder according to claim 28, wherein the correspondence information is stored in a header area of the bitstream.

30. The encoder according to claim 16, wherein the identifier is stored in a header area of the image associated with the identifier.

31. The encoder according to claim 16, wherein

the image includes a plurality of images,

the identifier includes one or more identifiers, and

the one or more identifiers are stored in a header area of one image included in the plurality of images, and are associated with the one image and one or more images different from the one image.

32. The encoder according to any one of claims 29 to 31, wherein the header area is an SEI area.

33. The encoder according to claim 16, wherein

the image includes a plurality of images, and

the circuitry encodes, to the bitstream, specification information for specifying the image associated with the identifier among the plurality of images.

34. A decoding method comprising:

by a decoder,

decoding, from a bitstream, an image and an identifier associated with the image, the identifier indicating at least one of (i) the image being suitable for task processing corresponding to the identifier and (ii) the image being not suitable for the task processing.

35. An encoding method comprising:

by an encoder,

encoding, to a bitstream, an image and an identifier associated with the image, the identifier indicating at least one of (i) the image being suitable for task processing corresponding to the identifier and (ii) the image being not suitable for the task processing.