as an enhancement data unit, data in a second frequency range higher than the first frequency range of the face image. The enhancement data unit is one of the one or more enhancement data units.

[0071]

With this, it may be possible to encode the low-frequency component data of the face image as the base data unit, and encodes the high-frequency component data of the face image as the enhancement data unit. In the reconstruction of the face video, it may be possible to apply the low-frequency component data to generate the first frame of the face video, and apply both the low-frequency component data and the high-frequency component data to generate the second frame of the face video. Accordingly, it may be possible to cause less discomfort in the face video while reducing delay.

[0072]

Moreover, an encoder of Example 28 may be the encoder of any one of Examples 19 to 25, in which the circuitry encodes, as the base data unit, a first image that (i) is related to the face image and (ii) has a first resolution, and encodes, as an enhancement data unit, a second image that (i) is related to the face image, (ii) is encoded using the first image as reference, and (iii) has a second resolution higher than the first resolution. The enhancement data unit is one of the one or more enhancement data units.

[0073]

With this, it may be possible to encode the face image with low resolution as the base data unit, and encodes the face image with high resolution as the enhancement data unit. In the reconstruction of the face video, it may be possible to apply the face image with low resolution to generate the first frame of the face video, and apply the face image with high resolution to generate the second frame of the face video. Accordingly, it may be possible to cause less discomfort in the face video while reducing delay.

[0074]

Moreover, a decoding method of Example 29 includes: decoding, from a bitstream, a base data unit of a face image related to a face video and one or more enhancement data units of the face image; decoding, from the bitstream, geometric information corresponding to each of frames of the face video and indicating geometric attributes within a region including a face of a person; and generating the face video from the base data unit, the one or more enhancement data units, and the geometric information, using a generative

[0336]

With this, it may be possible to encode the low-frequency component data of the face image as the base data unit, and encodes the high-frequency component data of the face image as the enhancement data unit. In the reconstruction of the face video, it may be possible to apply the low-frequency component data to generate the first frame of the face video, and apply both the low-frequency component data and the high-frequency component data to generate the second frame of the face video. Accordingly, it may be possible to cause less discomfort in the face video while reducing delay.

[0337]

It is to be noted that the data in the first frequency range may include the lowest-frequency component data, i.e., direct-current (DC) component data. Moreover, the one or more enhancement data units may be one or more data sets in one or more second frequency ranges higher than the first frequency range. When the enhancement data units are used, the second frequency ranges different from each other may be used.

[0338]

Moreover, for example, circuitry 151 may encode, as the base data unit, a first image. Moreover, for example, circuitry 151 may encode, as an enhancement data unit, a second image. Here, the first image is related to the face image and has a first resolution. The second image is related to the face image, is encoded using the first image as reference, and has a second resolution higher than the first resolution. The enhancement data unit is one of the one or more enhancement data units.

[0339]

With this, it may be possible to encode the face image with low resolution as the base data unit, and encodes the face image with high resolution as the enhancement data unit. In the reconstruction of the face video, it may be possible to apply the face image with low resolution to generate the first frame of the face video, and apply the face image with high resolution to generate the second frame of the face video. Accordingly, it may be possible to cause less discomfort in the face video while reducing delay.

[0340]

It is to be noted that the one or more enhancement data units may correspond to one or more second images having one or more second resolutions higher than the first resolution. When the enhancement data units are used,

[CLAIMS]

[Claim 1]

A decoder comprising:

memory; and

circuitry coupled to the memory, wherein

using the memory, the circuitry:

decodes, from a bitstream, a base data unit of a face image related to a face video and one or more enhancement data units of the face image;

decodes, from the bitstream, geometric information corresponding to each of frames of the face video and indicating geometric attributes within a region including a face of a person; and

generates the face video from the base data unit, the one or more enhancement data units, and the geometric information, using a generative model,

in the bitstream, the base data unit is added to a data set corresponding to a first frame that is a frame of the face video, and

in the bitstream, the one or more enhancement data units are added to one or more data sets corresponding to one or more second frames that are one or more frames of the face video and follow the first frame.

[Claim 2]

The decoder according to claim 1, wherein

the circuitry decodes, from a header, control information regarding a control of at least one of face image data units that are the base data unit and the one or more enhancement data units.

[Claim 3]

The decoder according to claim 2, wherein

the control information includes presence information indicating whether a face image data unit is included in an access unit controlled by the header, the face image data unit being one of the face image data units.

[Claim 4]

The decoder according to claim 2 or 3, wherein

when a face image data unit is included in an access unit controlled by the header, the control information includes type information regarding whether the face image data unit is the base data unit or an enhancement data unit, the face image data unit being one of the face image data units, the enhancement data unit being one of the one or more enhancement data units,

when the access unit includes the base data unit, the type information indicates that the face image data unit included in the access unit is the base data unit and continues to be used until a next base data unit, and

when the access unit includes the enhancement data unit, the type information indicates that the face image data unit included in the access unit is the enhancement data unit and is used together with the base data unit.

[Claim 5]

The decoder according to claim 2 or 3, wherein

when a face image data unit is included in an access unit controlled by the header, the control information includes application information indicating whether the face image data unit is applicable to generate and display a frame corresponding to the access unit among the frames of the face video, the face image data unit being one of the face image data units.

[Claim 6]

The decoder according to any one of claims 1 to 3, wherein

each of the base data unit and the one or more enhancement data units is represented by a vector indicating a facial feature included in the face image.

[Claim 7]

The decoder according to any one of claims 1 to 3, wherein

each of the base data unit and the one or more enhancement data units is represented by an image related to the face image.

[Claim 8]

The decoder according to any one of claims 1 to 3, wherein

the circuitry inputs the base data unit, at least one of the one or more enhancement data units, and the geometric information to the generative model to generate a frame of the face video.

[Claim 9]

The decoder according to any one of claims 1 to 3, wherein

the circuitry generates an intermediate image from the base data unit and at least one of the one or more enhancement data units, and inputs the intermediate image and the geometric information to the generative model to generate a frame of the face video.

[Claim 10]

The decoder according to any one of claims 1 to 3, wherein

the circuitry decodes an enhancement data unit using the base data unit as reference, and inputs the enhancement data unit and the geometric information to the generative model to generate a frame of the face video, the enhancement data unit being one of the one or more enhancement data units.

[Claim 11]

The decoder according to any one of claims 1 to 3, wherein

the base data unit is data of part of a face included in the face image, and

an enhancement data unit is data of other part of the face included in the face image, the enhancement data unit being one of the one or more enhancement data units.

[Claim 12]

The decoder according to any one of claims 1 to 3, wherein

the base data unit is data in a first frequency range of the face image, and

an enhancement data unit is data in a second frequency range higher than the first frequency range of the face image, the enhancement data unit being one of the one or more enhancement data units.

[Claim 13]

The decoder according to any one of claims 1 to 3, wherein

the base data unit corresponds to a first image that (i) is related to the face image and (ii) has a first resolution, and

an enhancement data unit corresponds to a second image that (i) is related to the face image, (ii) is decoded using the first image as reference, and (iii) has a second resolution higher than the first resolution, the enhancement data unit being one of the one or more enhancement data units.

[Claim 14]

The decoder according to any one of claims 1 to 3, wherein

the base data unit corresponds to a first image that (i) is related to the face image and (ii) is decoded with a first quantization step size, and

an enhancement data unit corresponds to a second image that (i) is related to the face image and (ii) is decoded with a second quantization step size finer than the first quantization step size using the first image as reference, the enhancement data unit being one of the one or more enhancement data units.

[Claim 15]

The decoder according to claim 2 or 3, wherein

the control information includes identification information for identifying each of the one or more enhancement data units.

[Claim 16]

The decoder according to claim 2 or 3, wherein

the control information includes total number information (i) included in the header of an access unit including the base data unit and (ii) indicating a total number of the one or more enhancement data units.

[Claim 17]

The decoder according to claim 2 or 3, wherein

the control information includes specification information (i) included in the header of an access unit including the base data unit and (ii) for specifying an enhancement data unit that is applicable to generate and display a second frame corresponding to an access unit including the enhancement data unit, the enhancement data unit being among the one or more enhancement data units, the second frame being among the one or more second frames.

[Claim 18]

The decoder according to any one of claims 1 to 3, wherein

the circuitry decodes at least one control parameter for controlling a stream buffer at which the bitstream is stored in the memory, the at least one control parameter being for controlling a buffer size of the stream buffer to be smaller than or equal to a reference size and an initial delay time at start of a decoding process to be shorter than or equal to a reference delay time.

[Claim 19]

An encoder comprising:

memory; and

circuitry coupled to the memory, wherein

using the memory, the circuitry:

encodes, into a bitstream, a base data unit of a face image related to a face video and one or more enhancement data units of the face image; and

encodes, into the bitstream, geometric information corresponding to each of frames of the face video and indicating geometric attributes within a region including a face of a person,

in the bitstream, the base data unit is added to a data set corresponding to a first frame that is a frame of the face video, and

in the bitstream, the one or more enhancement data units are added to one or more data sets corresponding to one or more second frames that are one or more frames of the face video and follow the first frame.

[Claim 20]

The encoder according to claim 19, wherein

the circuitry encodes, into a header, control information regarding a control of at least one of face image data units that are the base data unit and the one or more enhancement data units.

[Claim 21]

The encoder according to claim 20, wherein

the control information includes presence information indicating whether a face image data unit is included in an access unit controlled by the header, the face image data unit being one of the face image data units.

[Claim 22]

The encoder according to claim 20 or 21, wherein

when a face image data unit is included in an access unit controlled by the header, the control information includes type information regarding whether the face image data unit is the base data unit or an enhancement data unit, the face image data unit being one of the face image data units, the enhancement data unit being one of the one or more enhancement data units,

when the access unit includes the base data unit, the type information indicates that the face image data unit included in the access unit is the base data unit and continues to be used until a next base data unit, and

when the access unit includes the enhancement data unit, the type information indicates that the face image data unit included in the access unit is the enhancement data unit and is used together with the base data unit.

[Claim 23]

The encoder according to claim 20 or 21, wherein

when a face image data unit is included in an access unit controlled by the header, the control information includes application information indicating whether the face image data unit is applicable to generate and display a frame corresponding to the access unit among the frames of the face video, the face image data unit being one of the face image data units.

[Claim 24]

The encoder according to any one of claims 19 to 21, wherein

each of the base data unit and the one or more enhancement data units is represented by a vector indicating a facial feature included in the face image.

[Claim 25]

The encoder according to any one of claims 19 to 21, wherein

each of the base data unit and the one or more enhancement data units is represented by an image related to the face image.

[Claim 26]

The encoder according to any one of claims 19 to 21, wherein

the circuitry derives and encodes, as the base data unit, data of part of a face included in the face image, and derives and encodes, as an enhancement data unit, data of other part of the face included in the face image, the enhancement data unit being one of the one or more enhancement data units.

[Claim 27]

The encoder according to any one of claims 19 to 21, wherein

the circuitry derives and encodes, as the base data unit, data in a first frequency range of the face image, and derives and encodes, as an enhancement data unit, data in a second frequency range higher than the first frequency range of the face image, the enhancement data unit being one of the one or more enhancement data units.

[Claim 28]

The encoder according to any one of claims 19 to 21, wherein

the circuitry encodes, as the base data unit, a first image that (i) is related to the face image and (ii) has a first resolution, and encodes, as an enhancement data unit, a second image that (i) is related to the face image, (ii) is encoded using the first image as reference, and (iii) has a second resolution higher than the first resolution, the enhancement data unit being one of the one or more enhancement data units.

[Claim 29]

A decoding method comprising:

decoding, from a bitstream, a base data unit of a face image related to a face video and one or more enhancement data units of the face image;

decoding, from the bitstream, geometric information corresponding to each of frames of the face video and indicating geometric attributes within a region including a face of a person; and

generating the face video from the base data unit, the one or more enhancement data units, and the geometric information, using a generative model, wherein

in the bitstream, the base data unit is added to a data set corresponding to a first frame that is a frame of the face video, and

in the bitstream, the one or more enhancement data units are added to one or more data sets corresponding to one or more second frames that are one or more frames of the face video and follow the first frame.

[Claim 30]

An encoding method comprising:

encoding, into a bitstream, a base data unit of a face image related to a face video and one or more enhancement data units of the face image; and

encoding, into the bitstream, geometric information corresponding to each of frames of the face video and indicating geometric attributes within a region including a face of a person, wherein

in the bitstream, the base data unit is added to a data set corresponding to a first frame that is a frame of the face video, and

in the bitstream, the one or more enhancement data units are added to one or more data sets corresponding to one or more second frames that are one or more frames of the face video and follow the first frame.