CLAIMS

1. A method for video compression processing, executed by a processor of an electronic device, the method comprising:

performing video compression by using an end-to-end model trained with a single loss function, to obtain video compression information, comprising:

obtaining first motion information according to a current original frame and a previous reconstruction frame;

obtaining a current prediction frame according to the first motion information and the previous reconstruction frame; and

obtaining first residual information according to the current original frame and the current prediction frame, the first motion information and the first residual information forming the video compression information; and

wherein a loss value of the end-to-end model is determined with the single loss function, based on the first motion information, the first residual information, the current original frame, and the current prediction frame.

2. The method of claim 1, wherein performing video compression by using the end-to-end model trained with the single loss function, to obtain the video compression information, further comprises:

obtaining a current reconstruction frame according to the first residual information and the current prediction frame, the current reconstruction frame and a next original frame being used jointly to obtain the first motion information for compression of the next original frame.

3. The method of claim 1, wherein obtaining the current prediction frame according to the first motion information and the previous reconstruction frame comprises:

obtaining a motion feature based on the first motion information;

obtaining second motion information by decoding based on the motion feature; and

obtaining the current prediction frame based on the previous reconstruction frame and the second motion information.

4. The method of claim 3, wherein obtaining the motion feature based on the first motion information comprises:

encoding the first motion information to obtain a first motion feature;

wherein the method further comprises:

quantizing the first motion feature to obtain a second motion feature; and

wherein obtaining the second motion information by decoding based on the motion feature comprises:

decoding the second motion feature to obtain the second motion information.

5. The method of claim 4, wherein encoding the first motion information to obtain the first motion feature comprises:

convoluting, by a first convolutional layer, the first motion information to obtain a first encoded value;

normalizing, by a first normalization layer, the first encoded value, to obtain a first normalized value;

convoluting, by an n^{th} convolutional layer, an $(n-1)^{th}$ normalized value output by an $(n-1)^{th}$ normalization layer, to obtain an n^{th} encoded value;

normalizing, by an n^{th} normalization layer, the n^{th} encoded value, to obtain an n^{th} normalized value; and

normalizing, by an N^{th} convolutional layer, an $(N-1)^{th}$ normalized value to obtain the first motion feature,

where n is a positive integer greater than 1 and smaller than N, and N denotes a number of convolutional layers.

6. The method of claim 5, wherein decoding the second motion feature to obtain the second motion information comprises:

deconvoluting, by a first deconvolutional layer, the first motion feature, to obtain a first decoded value;

denormalizing, by a first denormalization layer, the first decoded value, to obtain a first denormalized value;

deconvoluting, by an n^{th} deconvolutional layer, an $(n-1)^{th}$ denormalized value output by an $(n-1)^{th}$ denormalization layer, to obtain an n^{th} decoded value;

normalizing, by an n^{th} denormalization layer, the n^{th} decoded value, to output an n^{th} denormalized value; and

deconvoluting, by an N^{th} deconvolutional layer, an $(N-1)^{th}$ denormalized value, to obtain the second motion information,

where n is a positive integer greater than 1 and smaller than the N, and N denotes a number of deconvolutional layers.

7. The method of any one of claims 1, further comprising:

encoding the first residual information to obtain a first residual feature; and

quantizing the first residual feature to obtain a second residual feature,

wherein the second residual feature and a second motion feature formed based on the first motion information are constituent parts of the video compression information.

8. The method of claim 3, wherein obtaining the current prediction frame based on the previous reconstruction frame and the second motion information comprises:

splicing the previous reconstruction frame, the second motion information and a projection frame to obtain a first splicing frame, the projection frame being generated based on convolution of the previous reconstruction frame and the second motion information;

convoluting the first splicing frame to obtain a first convolutional frame;

performing (S+1) times of first residual processing and S times of pooling processing based on the first convolutional frame, to obtain a first intermediate frame, the first residual processing and the pooling processing being performed alternately;

performing (S+1) times of second residual processing and S times of upsampling processing on the first intermediate frame, to obtain a second intermediate frame, the second residual processing and the upsampling processing being performed alternately;

performing second convolutional processing on the second intermediate frame to obtain a second convolutional frame; and

performing third convolutional processing on the second convolutional frame to obtain the current prediction frame,

wherein the upsampling processing is performed in a bilinear interpolation manner.

9. An electronic device, comprising:

a memory having computer executable instructions stored therein; and

a processor, connected to the memory, and configured, when executing computer executable instructions stored in the memory, to:

perform video compression by using an end-to-end model trained with a single

loss function, to obtain video compression information, wherein the processor is configured to:

obtain first motion information according to a current original frame and a previous reconstruction frame;

obtain a current prediction frame according to the first motion information and the previous reconstruction frame; and

obtain first residual information according to the current original frame and the current prediction frame, the first motion information and the first residual information forming the video compression information,

wherein a loss value of the end-to-end model is determined with the single loss function, based on the first motion information, the first residual information, the current original frame, and the current prediction frame.

10. The electronic device of claim 9, wherein the processor is further configured to:

obtain a current reconstruction frame according to the first residual information and the current prediction frame, the current reconstruction frame and a next original frame being used jointly to obtain the first motion information for compression of the next original frame.

11. The electronic device of claim 9, wherein the processor is configured to:

encode the first motion information to obtain a first motion feature;

quantize the first motion feature to obtain a second motion feature;

decode the second motion feature to obtain second motion information; and

obtain the current prediction frame based on the previous reconstruction frame and the second motion information.

12. The electronic device of claim 11, wherein the processor is configured to:

convolute, at a first convolutional layer, the first motion information to obtain a first encoded value;

normalize, at a first normalization layer, the first encoded value to obtain a first normalized value;

convolute, at an n^{th} convolutional layer, an $(n-1)^{th}$ normalized value output by an $(n-1)^{th}$ normalization layer to obtain an n^{th} encoded value;

normalize, at an n^{th} normalization layer, the n^{th} encoded value to obtain an n^{th} normalized value; and

normalize, at an N^{th} convolutional layer, an $(N-1)^{th}$ normalized value to obtain the first motion feature ,

where n is a positive integer greater than 1 and smaller than N, and N denotes a number of convolutional layers.

13. The electronic device of claim 11, wherein the processor is configured to:

deconvolute, at a first deconvolutional layer, the first motion feature to obtain a first decoded value;

denormalize, at a first denormalization layer, the first decoded value to obtain a first denormalized value;

deconvolute, at an n^{th} deconvolutional layer, an $(n-1)^{th}$ denormalized value output by an $(n-1)^{th}$ denormalization layer to obtain an n^{th} decoded value;

normalize, at an n^{th} denormalization layer, the n^{th} decoded value to output an n^{th} denormalized value; and

deconvolute, at an N^{th} deconvolutional layer, an $(N-1)^{th}$ denormalized value to obtain

the second motion information,

where n is a positive integer greater than 1 and smaller than N, and N denotes a number of deconvolutional layers.

14. The electronic device of claim 9, wherein the processor is further configured to:

encode the first residual information to obtain a first residual feature; and

quantize the first residual feature to obtain a second residual feature,

wherein the second residual feature and the second motion feature formed based on the first motion information are constituent parts of the video compression information.

15. The electronic device of claim 11, wherein the processor is configured to:

splice the previous reconstruction frame, the second motion information and a projection frame to obtain a first splicing frame, the projection frame being generated based on convolution of the previous reconstruction frame and the second motion information; and

convolute the first splicing frame to obtain a first convolutional frame;

perform (S+1) times of first residual processing and S times of pooling processing based on the first convolutional frame to obtain a first intermediate frame, the first residual processing and the pooling processing being performed alternately;

perform (S+1) times of second residual processing and S times of upsampling processing on the first intermediate frame to obtain a second intermediate frame, the second residual processing and the upsampling processing being performed alternately;

perform second convolutional processing on the second intermediate frame to obtain a second convolutional frame; and

perform third convolutional processing on the second convolutional frame to obtain the current prediction frame,

wherein the upsampling processing is performed in a bilinear interpolation manner.

16. A non-transitory computer storage medium, storing computer executable instructions, which, executed by a processor , cause the processor to implement the method of any one of claims 1 to 8.