

DESCRIPTION

TITLE OF THE INVENTION

DYNAMIC IMAGE PREDICTIVE ENCODING AND DECODING
DEVICE, METHOD, AND PROGRAM

5 **Technical Field**

[0001] The present invention relates to a video predictive encoding device, method, and program and a video predictive decoding device, method, and program and, more particularly, to a video predictive encoding device, method, and program and a video predictive decoding device, method, and program associated with inter-frame prediction effective to random access.

Background Art

[0002] Compression techniques are used for efficient transmission and storage of video data. The techniques according to MPEG1-4 and H.261-H.264 are widely used for compressing video data.

[0003] In these compressing techniques, a target picture to be encoded is partitioned into a plurality of blocks which are then subjected to encoding and decoding. The predictive encoding methods as described below are used for enhancement of encoding efficiency. In intra-frame predictive encoding, a predicted signal is generated using a decoded neighboring picture signal (a decoded signal from picture data compressed in the past) present in the same frame as a target block, and then a difference signal obtained by subtracting the predicted signal from a signal of the target block is encoded. In inter-frame predictive encoding, a displacement of signal is searched for with reference to a reconstructed picture signal present in a frame different from a target

block, a predicted signal is generated with compensation for the displacement, and a difference signal obtained by subtracting the predicted signal from the signal of the target block is encoded. The reconstructed picture used for reference for the motion search and compensation is referred to as a reference picture.

[0004] In bidirectional inter-frame prediction, reference can be made not only to past pictures in the output time order, but also to future pictures following the target picture in the output time order (provided that the future pictures are encoded prior to the target picture and preliminarily reproduced). A predicted signal derived from a past picture and a predicted signal derived from a future picture can be averaged, to provide for effective prediction of a newly-appearing object in a picture, and to reduce noise included in the two predicted signals.

[0005] Furthermore, in the inter-frame predictive encoding of H.264, the predicted signal for the target block is selected by performing the motion search with reference to a plurality of reference pictures which have previously been encoded and then reproduced, and by defining a picture signal with the smallest error as an optimum predicted signal. A difference is calculated between the pixel signal of the target block and this optimum predicted signal, which is then subjected to a discrete cosine transform, quantization, and entropy encoding. At the same time, information regarding a reference picture and a region from which the optimum predicted signal for the target block is derived (which will be respectively referred to as "reference index" and "motion vector") are also encoded. In H.264, four or five reproduced

pictures are stored as reference pictures in a frame memory or decoded picture buffer.

[0006] The inter-frame predictive encoding allows efficient compression encoding by taking advantage of correlation between pictures, however dependence between frames is avoided in order to allow viewing of a video program from the middle, such as when switching between TV channels. Points having no dependence between frames in a compressed bit stream of a video sequence are referred to as "random access points." Besides the switching of channels, the random access points can also be used in cases of editing a video sequence and joining compressed data of different video sequences. In the conventional technology, "clean random access points" are provided as random access points. The clean random access points are specified by clean random access pictures (which will be referred to hereinafter as "CRA pictures") of Network Abstraction Layer (NAL) unit type. One bit stream can include a plurality of CRA pictures and a video predictive decoding device may start decoding from any clean random access point.

[0007] In the described embodiments, picture types of pictures associated with a CRA picture are defined as follows (cf. Fig. 10).

a) Past picture: picture decoded before the CRA picture and preceding the CRA picture in output order.

b) Lagging picture: picture decoded before the CRA picture but following the CRA picture in output order.

c) Leading picture: picture decoded after the CRA picture but preceding the CRA picture in output order.

d) Normal picture: picture decoded after the CRA picture and following the CRA picture in output order.

[0008] Since the CRA picture is defined as a picture limited only to intra-frame prediction, it is provided with all information necessary for decoding and can be correctly decoded without reference to any other picture. Every normal picture following the CRA picture is defined so that inter-frame prediction from a past picture, a lagging picture, or a leading picture is prohibited.

[0009] Decoding of a bit stream from a CRA picture and normal pictures are correctly decoded without errors in inter-frame prediction. However, leading pictures which are decoded after the CRA picture may, or may not be correctly decoded without errors in inter-frame prediction. In other words, there are correctly-decoded leading pictures, while there can also be incorrectly-decoded leading pictures.

[0010] The term "correctly-decoded" herein means that a decoded picture is the same as a picture obtained in an operation of decoding a bit stream not from the CRA picture, but instead from the head of the bit stream. In decoding from a CRA picture, a picture (e.g., a lagging picture) preceding the CRA picture in decoding order is not decoded and it does not exist in the decoded picture buffer. Therefore, a subsequent picture the inter-frame prediction of which is carried out directly or indirectly using a picture preceding the CRA picture in decoding order can include a decoding error.

Citation List

Non Patent Literature

[0011] Non Patent Literature 1: Benjamin Bross et al., "High efficiency

video coding (HEVC) text specification draft 7," Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, 9th Meeting: Geneva, CH, 27 April - 7 May 2012

SUMMARY OF THE INVENTION

5 Technical Problem

[0012] As described above, when the video predictive decoding device starts decoding from a random access point, there is the possibility of existence of an incorrectly-decoded picture, and the incorrectly-decoded picture should not be used for decoding thereafter.

10 On the other hand, in the case where there is a correctly-decoded picture, the correctly-decoded picture can be used for decoding thereafter. Since the conventional technologies have no method for specifying which picture following the random access point in decoding order should be discarded, all leading pictures are handled as
15 pictures that cannot be correctly decoded, and are thus discarded. However, some of these pictures can in fact be decoded, and can contribute to improvement in prediction performance of subsequent pictures; therefore, discarding all of the leading pictures as incorrectly-decoded pictures is not desirable.

20 [0013] It is therefore an object of the present invention to enable identification of a decodable picture so as to make the decodable picture available as a reference picture for a subsequent picture, thereby contributing to improvement in prediction performance.

Solution to Problem

25 [0014] A video predictive encoding device according to an embodiment of the present invention comprises: input means which

inputs a plurality of pictures constituting a video sequence; encoding means which encodes the pictures by a method of either intra-frame prediction or inter-frame prediction to generate compressed picture data and which also encodes output order information of each picture and information about a picture type of each picture, the compressed picture data generated to include a picture serving as a random access point;; reconstruction means which decodes the compressed picture data to reconstruct pictures; picture storage means which stores one or more of the reconstructed pictures as reference pictures to be used for encoding of a subsequent picture; and control means which determines the picture type and controls the picture storage means, based on the determination of the picture type, wherein the control means labels each of the pictures as one of three types defined below: 1) a CRA picture: a picture which is characterized in that a type 2 picture subsequent to a CRA picture can be correctly decoded when decoding is started from the CRA picture; 2) a type 1 picture: a picture which is decoded after a CRA picture associated with the picture, and is outputted before the associated CRA picture, which is not subjected to a list of reference pictures including at least one reference picture labeled as a type 1 picture or at least one reference picture preceding said associated CRA picture in decoding order, for execution of inter-frame prediction; 3) a type 2 picture: a picture which has a list of reference pictures, for execution of inter-frame prediction, and which is characterized in that every reference picture in the list of reference pictures is labeled as either a type 2 picture or a CRA picture and is decoded after a CRA picture associated with the picture.

[0015] A video predictive encoding method according to an embodiment of the present invention is a video predictive encoding method executed by a video predictive encoding device, comprising: an input step of inputting a plurality of pictures constituting a video sequence; an encoding step of encoding the pictures by a method of either intra-frame prediction or inter-frame prediction to generate compressed picture data including a picture serving as a random access point and also encoding output order information of each picture and information about a picture type of each picture; a reconstruction step of decoding the compressed picture data to reconstruct pictures; a picture storage step of storing one or more of the reconstructed pictures as reference pictures to be used for encoding of a subsequent picture; and a control step of determining the picture type and controlling the picture storage step, based on the determined picture type, wherein the control step further comprises the video predictive encoding device labeling each of the pictures as one of three types defined below: 1) CRA picture: a picture which is characterized in that a type 2 picture, subsequent to a CRA picture, can be correctly decoded when decoding is started from the CRA picture; 2) type 1 picture: a picture which is decoded after a CRA picture associated with the picture and is outputted before the associated CRA picture, and which has a list of reference pictures for execution of inter-frame prediction, the list of reference pictures including at least one reference picture labeled as a type 1 picture, or at least one reference picture preceding said associated CRA picture in decoding order; 3) type 2 picture: a picture which has a list of reference pictures, for execution of inter-frame

prediction, and which is characterized in that every reference picture in the list of reference pictures is labeled as a type 2 picture or as a CRA picture and is decoded after an associated CRA picture.

[0016] A video predictive encoding program according to the present

invention is a video predictive encoding program for letting a computer function as: input means that inputs a plurality of pictures constituting a video sequence; encoding means which encodes the pictures by a method of either intra-frame prediction or inter-frame prediction to generate compressed picture data including a picture serving as a

random access point and which also encodes output order information of each picture and information about a picture type of each picture;

reconstruction means which decodes the compressed picture data to reconstruct pictures; picture storage means which stores one or more of the reconstructed pictures as reference pictures to be used for encoding

of a subsequent picture; and control means which determines the picture type and controls the picture storage means, based on the determination result, wherein the control means labels each of the pictures as one of three types defined below: 1) CRA picture: a picture which is characterized in that a type 2 picture subsequent to a CRA

picture can be correctly decoded when decoding is started from the CRA picture; 2) type 1 picture: a picture which is decoded after a CRA

picture associated with the picture, and is outputted before the associated CRA picture, and which has a list of reference pictures for

execution of inter-frame prediction, the list of reference pictures including at least one reference picture labeled as a type 1 picture or at

least one reference picture preceding said associated CRA picture in

decoding order;; 3) type 2 picture: a picture which has a list of reference pictures, for execution of inter-frame prediction, and which is characterized in that every reference picture in the list of reference pictures is labeled as a type 2 picture or as a CRA picture and decoded after a CRA picture associated with the picture.

[0017] A video predictive decoding device according to an embodiment of the present invention is a video predictive decoding device comprising: input means that inputs, for a plurality of pictures constituting a video sequence, compressed picture data including a random access picture and encoded data indicative of an output order of each picture and a picture type of each picture, resulting from encoding by either intra-frame prediction or inter-frame prediction; reconstruction means which decodes the compressed picture data and the encoded data to reconstruct pictures, output order information, and output picture type information; picture storage means which stores one or more of said reconstructed pictures as reference pictures to be used for decoding of a subsequent picture; and control means which controls the reconstruction means, based on the picture type, wherein each picture is labeled with the picture type as one of three types defined below: 1) CRA picture: a picture which is characterized in that a type 2 picture subsequent to a CRA picture, can be correctly decoded when decoding is started from the CRA picture; 2) type 1 picture: a picture which is decoded after an associated CRA picture and is outputted before the associated CRA picture, and which has a list of reference pictures for execution of inter-frame prediction, the list of reference pictures including at least one reference picture labeled as a type 1

picture or at least one reference picture preceding said associated CRA picture in decoding order; 3) type 2 picture: a picture which has a list of reference pictures, for execution of inter-frame prediction, and which is characterized in that every reference picture in the list of reference pictures is labeled as a type 2 picture or as a CRA picture and decoded after an associated CRA picture; and wherein the reconstruction means continues, during a period immediately before a process of a next CRA picture, a decoding process such that when decoding of encoded data is started from a CRA picture, the reconstruction means decodes a picture labeled as a type 2 picture and skips decoding of a picture labeled as a type 1 picture.

[0018] A video predictive decoding method according to an embodiment of the present invention is a video predictive decoding method executed by a video predictive decoding device, comprising: an input step of inputting compressed picture data including a random access picture and encoded data indicative of an output order of each picture and a picture type of each picture, resulting from encoding by either inter-frame prediction or intra-frame prediction for a plurality of pictures constituting a video sequence; a reconstruction step of decoding the compressed picture data and the encoded data to reconstruct pictures, output order information, and picture type information; a picture storage step of storing one or more of said reconstructed pictures as reference pictures to be used for decoding of a subsequent picture; and a control step of controlling the reconstruction step, based on the picture type, wherein each picture is labeled with the picture type as one of three types defined below: 1) CRA picture: a

picture which is characterized in that a type 2 picture subsequent to a CRA picture can be correctly decoded when decoding is started from the CRA picture; 2) type 1 picture: a picture which is decoded after a CRA picture associated with the picture, and is outputted before the associated CRA picture, and which has a list of reference pictures for execution of inter-frame prediction, the list of reference pictures including at least one reference picture labeled as a type 1 picture or at least one reference picture preceding said associated CRA picture in decoding order; 3) type 2 picture: a picture which has a list of reference pictures, for execution of inter-frame prediction, and which is characterized in that every reference picture in the list of reference pictures is labeled as either a type 2 picture or a CRA picture and is decoded after a CRA picture associated with the picture; and wherein in the reconstruction step the video predictive decoding device continues, during a period immediately before a process of a next CRA picture, a decoding process such that when decoding of encoded data is started from a CRA picture, the video predictive decoding device decodes a picture labeled as a type 2 picture and skips decoding of a picture labeled as a type 1 picture.

[0019] A video predictive decoding program according to an embodiment of the present invention is a video predictive decoding program for letting a computer function as: input means that inputs compressed picture data including a random access picture and encoded data indicative of an output order of each picture and a picture type of each picture, resulting from encoding by either intra-frame prediction or inter-frame prediction for a plurality of pictures

constituting a video sequence; reconstruction means which decodes the compressed picture data and the encoded data to reconstruct pictures, output order information, and picture type information; picture storage means which stores one or more of said reconstructed pictures as reference pictures to be used for decoding of a subsequent picture; and control means which controls the reconstruction means, based on the picture type, wherein each picture is labeled with the picture type as one of three types defined below: 1) CRA picture: a picture which is characterized in that a type 2 picture subsequent to a CRA picture can be correctly decoded when decoding is started from the CRA picture; 2) type 1 picture: a picture which is decoded after a CRA picture that is associated with the picture, and is outputted before the associated CRA picture, and which has a list of reference pictures for execution of inter-frame prediction, the list of reference pictures including at least one reference picture labeled as a type 1 picture or at least one reference picture preceding said associated CRA picture in decoding order; 3) type 2 picture: a picture which has a list of reference pictures, for execution of inter-frame prediction, and which is characterized in that every reference picture in the list of reference pictures is labeled as either a type 2 picture or a CRA picture and is decoded after a CRA picture associated with the picture; and wherein the reconstruction means continues, during a period immediately before a process of a next CRA picture, a decoding process such that when decoding of encoded data is started from a CRA picture, the reconstruction means decodes a picture labeled as a type 2 picture and skips decoding of a picture labeled as a type 1 picture.

[0020] It should be noted herein that the video predictive encoding device, method, and program and the video predictive decoding device, method, and program according to embodiments of the present invention can also be realized employing the modes as described below.

[0021] Another video predictive encoding device according to an embodiment of the present invention is a video predictive encoding device comprising: input means that inputs a plurality of pictures constituting a video sequence; encoding means which encodes the pictures by a method of either intra-frame prediction or inter-frame prediction to generate compressed picture data including a picture serving as a random access point, and which also encodes output order information of each picture; reconstruction means which decodes the compressed picture data to reconstruct pictures; picture storage means which stores one or more of the reconstructed pictures as reference pictures to be used for encoding of a subsequent picture; and control means which controls the picture storage means, wherein the control means classifies and controls each of the pictures into three types defined below: 1) a CRA picture from which decoding of encoded data is started; 2) a picture which is decoded after a CRA picture associated with the picture, and is outputted before the associated CRA picture, which is not subjected to a decoding process by the reconstruction means and is not stored in the picture storage means or output, and which has a list of reference pictures for execution of inter-frame prediction, the list of reference pictures including at least one reference picture not subjected to the decoding process by the reconstruction

means, or at least one reference picture preceding the associated CRA picture in decoding order; 3) a picture which is decoded by the reconstruction means and stored in the picture storage means for reference as needed, and which is characterized in that the picture has a list of reference pictures for execution of inter-frame prediction and in that every reference picture in the list of reference pictures is decoded by the reconstruction means and is decoded after a CRA picture associated with the picture.

[0022] Another video predictive encoding method according to an embodiment of the present invention is a video predictive encoding method executed by a video predictive encoding device, comprising: an input step of inputting a plurality of pictures constituting a video sequence; an encoding step of encoding the pictures by a method of either intra-frame prediction or inter-frame prediction to generate compressed picture data including a picture serving as a random access point and also encoding output order information of each picture; a reconstruction step of decoding the compressed picture data to reconstruct pictures; a picture storage step of storing one or more of the reconstructed pictures as reference pictures to be used for encoding of a subsequent picture; and a control step of controlling the picture storage step, wherein in the control step the video predictive encoding device classifies and controls each of the pictures into three types defined below: 1) a CRA picture from which decoding of encoded data is started; 2) a picture which is decoded after a CRA picture associated with the picture, and is outputted before the associated CRA picture, which is not subjected to a decoding process by the reconstruction step

and is not stored in the picture storage step or output, and which has a list of reference pictures for execution of inter-frame prediction, the list of reference pictures including at least one reference picture not subjected to the decoding process by the reconstruction step, or at least one reference picture preceding the associated CRA picture in decoding order; 3) a picture which is decoded by the reconstruction step and stored in the picture storage step for reference as needed, and which is characterized in that the picture has a list of reference pictures for execution of inter-frame prediction and in that every reference picture in the list of reference pictures is decoded by the reconstruction step and is decoded after a CRA picture associated with the picture.

[0023] Another video predictive encoding program according to an embodiment of the present invention is a video predictive encoding program for letting a computer function as: input means that inputs a plurality of pictures constituting a video sequence; encoding means which encodes the pictures by a method of either intra-frame prediction or inter-frame prediction to generate compressed picture data including a picture serving as a random access point and which also encodes output order information of each picture; reconstruction means which decodes the compressed picture data to reconstruct pictures; picture storage means which stores one or more of the reconstructed pictures as reference pictures to be used for encoding of a subsequent picture; and control means which controls the picture storage means, wherein the control means classifies and controls each of the pictures into three types defined below: 1) a CRA picture from which decoding of encoded data is started; 2) a picture which is decoded after a CRA

picture associated with the picture, which is outputted after the associated CRA picture, which is not subjected to a decoding process by the reconstruction means and is not stored in the picture storage means or output, and which has a list of reference pictures for execution of inter-frame prediction, the list of reference pictures including at least one reference picture that is not subjected to the decoding process by the reconstruction means, or at least one reference picture preceding the associated CRA picture in decoding order; 3) a picture which is decoded by the reconstruction means and stored in the picture storage means for reference as needed, and which is characterized in that the picture has a list of reference pictures for execution of inter-frame prediction and in that every reference picture in the list of reference pictures is decoded by the reconstruction means and is decoded after a CRA picture associated with the picture.

[0024] Another video predictive decoding device according to an embodiment of the present invention is a video predictive decoding device comprising: input means that inputs compressed picture data including a random access picture and encoded data indicative of an output order of each picture, resulting from encoding by either intra-frame prediction or inter-frame prediction for a plurality of pictures constituting a video sequence; reconstruction means which decodes the compressed picture data and the encoded data to reconstruct pictures and output order information; picture storage means which stores one or more of said reconstructed pictures as reference pictures to be used for decoding of a subsequent picture; and control means which controls the reconstruction means, wherein the

control means classifies and controls each of the pictures into three types defined below: 1) a CRA picture from which decoding of encoded data is started; 2) a picture which is decoded after a CRA picture associated with the picture, and is outputted before the associated CRA picture, which is not subjected to a decoding process by the reconstruction means and is not stored in the picture storage means or output, and which has a list of reference pictures for execution of inter-frame prediction, the list of reference pictures including at least one reference picture which is not subjected to the decoding process by the reconstruction means, or at least one reference picture preceding the associated CRA picture in decoding order; 3) a picture which is decoded by the reconstruction means and stored in the picture storage means for reference as needed, and which is characterized in that the picture has a list of reference pictures for execution of inter-frame prediction and in that every reference picture in the list of reference pictures is decoded by the reconstruction means and decoded after a CRA picture associated with the picture; and wherein the reconstruction means continues, during a period immediately before a process of a next CRA picture, a decoding process such that when decoding of encoded data is started from a CRA picture associated with the picture, the reconstruction means determines whether every reference picture in a list of reference pictures for a target picture is stored in the picture storage means, that if every reference picture in the list of reference pictures is stored, the reconstruction means decodes the target picture, and that if one or more reference pictures in the list of reference pictures are not stored, the

reconstruction means skips decoding of the target picture.

[0025] Another video predictive decoding method according to an embodiment of the present invention is a video predictive decoding method executed by a video predictive decoding device, comprising: an
5 input step of inputting compressed picture data including a random access picture and encoded data indicative of an output order of each picture, resulting from encoding by either intra-frame prediction or inter-frame prediction for a plurality of pictures constituting a video sequence; a reconstruction step of decoding the compressed picture data and the encoded data to reconstruct pictures and output order
10 information; a picture storage step of storing one or more of said reconstructed pictures as reference pictures to be used for decoding of a subsequent picture; and a control step of controlling the reconstruction step, wherein in the control step the video predictive decoding device classifies and controls each of the pictures into three types defined
15 below: 1) a CRA picture from which decoding of encoded data is started; 2) a picture which is decoded after a CRA picture associated with the picture, and is outputted before the associated CRA picture, which is not subjected to a decoding process by the reconstruction step and is not stored in the picture storage step or output, and which has a
20 list of reference pictures for execution of inter-frame prediction, the list of reference pictures including at least one reference picture which is not subjected to the decoding process by the reconstruction step, or at least one reference picture preceding the associated CRA picture in decoding order; 3) a picture which is decoded by the reconstruction
25 step and is stored in the picture storage step for reference as needed,

and which is characterized in that the picture has a list of reference pictures for execution of inter-frame prediction and in that every reference picture in the list of reference pictures is decoded by the reconstruction step and is decoded after a CRA picture associated with the picture; wherein in the reconstruction step the video predictive decoding device continues, during a period immediately before a process of a next CRA picture, a decoding process such that when decoding of encoded data is started from an associated CRA picture, the video predictive decoding device determines whether every reference picture in a list of reference pictures for a target picture is stored in the picture storage step, that if every reference picture in the list of reference pictures is stored, the video predictive decoding device decodes the target picture, and that if one or more reference pictures in the list of reference pictures are not stored, the video predictive decoding device skips decoding of the target picture.

[0026] Another video predictive decoding program according to an embodiment of the present invention is a video predictive decoding program for letting a computer function as: input means that inputs compressed picture data including a random access picture and encoded data indicative of an output order of each picture, resulting from encoding by either intra-frame prediction or inter-frame prediction for a plurality of pictures constituting a video sequence; reconstruction means which decodes the compressed picture data and the encoded data to reconstruct pictures and output order information; picture storage means which stores one or more of said reconstructed pictures as reference pictures to be used for decoding of a subsequent

picture; and control means which controls the reconstruction means, wherein the control means classifies and controls each of the pictures into three types defined below: 1) a CRA picture from which decoding of encoded data is started; 2) a picture which is decoded after a CRA picture associated with the picture, and which is outputted after the associated CRA picture, which is not subjected to a decoding process by the reconstruction means and which is not stored in the picture storage means or output, and which has a list of reference pictures for execution of inter-frame prediction, the list of reference pictures including at least one reference picture which is not subjected to the decoding process by the reconstruction means, or at least one reference picture preceding the associated CRA picture in decoding order;; 3) a picture which is decoded by the reconstruction means and stored in the picture storage means for reference as needed, and which is characterized in that the picture has a list of reference pictures for execution of inter-frame prediction and in that every reference picture in the list of reference pictures is decoded by the reconstruction means and decoded after a CRA picture associated with the picture; and wherein the reconstruction means continues, during a period immediately before a process of a next CRA picture, a decoding process such that when decoding of encoded data is started from a CRA picture associated with the picture, the reconstruction means determines whether every reference picture in a list of reference pictures for a target picture is stored in the picture storage means, that if every reference picture in the list of reference pictures is stored, the reconstruction means decodes the target picture, and that if one or more

reference pictures in the list of reference pictures are not stored, the reconstruction means skips decoding of the target picture.

Effects of the Invention

[0027] Embodiments of the present invention enable discrimination of a decodable picture so as to make the decodable picture available as a reference picture for a subsequent picture, thereby contributing to improvement in prediction performance. More specifically, when decoding is started from a CRA picture at a leading end of a bit stream, the video predictive decoding device is able to detect whether a certain picture can be correctly decoded (by use of a label or by comparison with a reference picture set). For this reason, the video predictive decoding device can select and discard only a non-decodable picture (instead of discarding all leading pictures), so as to make a decodable picture available as a reference picture for a subsequent picture, thereby contributing to improvement in prediction performance.

BRIEF DESCRIPTION OF THE DRAWINGS

[0028] Fig. 1 is a block diagram showing a video predictive encoding device according to an embodiment of the present invention.

Fig. 2 is a block diagram showing a video predictive decoding device according to an embodiment of the present invention.

Fig. 3 is a drawing for explaining syntax elements according to an embodiment of the present invention.

Fig. 4 is a flowchart showing a video predictive encoding method according to an embodiment of the present invention.

Fig. 5 is a flowchart showing a video predictive decoding method according to an embodiment of the present invention.

Fig. 6 is a drawing showing a hardware configuration of a computer for executing a program stored in a storage medium.

Fig. 7 is a perspective view of a computer for executing a program stored in a storage medium.

5 Fig. 8 is a block diagram showing a configuration example of a video predictive encoding program.

Fig. 9 is a block diagram showing a configuration example of a video predictive decoding program.

10 Fig. 10 is a drawing for explaining the background of the present invention.

Embodiments of the Invention

[0029] Embodiments of the present invention will be described below using Figs. 1 to 9.

[0030] [About Video Predictive Encoding Device]

15 Fig. 1 is a function block diagram showing a configuration of a video predictive encoding device 100 according to an embodiment of the present invention. As shown in Fig. 1, the video predictive encoding device 100 is provided with an input terminal 101, a block divider 102, a predicted signal generator 103, a frame memory 104, a subtracter 105, a transformer 106, a quantizer 107, a de-quantizer 108,
20 an inverse-transformer 109, an adder 110, an entropy encoder 111, an output terminal 112, an input terminal 113, and a frame memory manager (or buffer manager) 114 as a functional configuration. The operations of the respective function blocks will be described in the operation of the video predictive encoding device 100 below. The
25 transformer 106 and quantizer 107 correspond to the encoding means

and the de-quantizer 108, inverse-transformer 109, and adder 110 correspond to the decoding means.

[0031] The operation of the video predictive encoding device 100 configured as described above will be described below. A video signal consisting of a plurality of pictures is fed to the input terminal 101. A picture of an encoding target is partitioned into a plurality of regions by the block divider 102. In the present embodiment, the target picture is partitioned into blocks each consisting of 8×8 pixels, but it may be partitioned into blocks of any size or shape other than the foregoing. A predicted signal is then generated for a region as a target of an encoding process (which will be referred to hereinafter as a target block). The present embodiment employs two types of prediction methods. Namely, they are the inter-frame prediction and the intra-frame prediction.

[0032] In the inter-frame prediction, reconstructed pictures which have been encoded and thereafter previously reconstructed are used as reference pictures, and motion information to provide the predicted signal with the smallest error from the target block is determined from the reference pictures. This process is called motion detection. Depending upon the situation, the target block can also be sub-divided into sub-regions to determine an inter-frame prediction method for each of the sub-regions. In this case, the most efficient division method for the entire target block and motion information of each sub-region are determined out of various division methods. In an embodiment according to the present invention, the operation is carried out in the predicted signal generator 103, the target block is fed via line L102,

and the reference pictures are fed via L104. The reference pictures to be used herein are a plurality of pictures which have been encoded and reconstructed in the past. The details of use of the reference pictures are the same as in the methods of MPEG-2 or 4 and H.264 which are the conventional technologies. Once the motion information and sub-region division method are determined as previously described, the motion information and sub-region division method are fed via line L112 to the entropy encoder 111 to be encoded thereby, and then the encoded data is output from the output terminal 112. Information indicating from which reference picture out of the plurality of reference pictures the predicted signal is derived (such information is called a "reference index") is also sent via line L112 to the entropy encoder 111. In an embodiment according to the present invention, four or five reconstructed pictures are stored in the frame memory 104 to be used as reference pictures. The predicted signal generator 103 derives reference picture signals from the frame memory 104, based on the reference pictures and motion information, corresponding to the sub-region division method and each sub-region, and generates the predicted signal. The inter-frame predicted signal generated in this manner is fed via line L103 to the subtracter 105.

[0033] In the intra-frame prediction, an intra-frame predicted signal is generated using reconstructed pixel values spatially adjacent to the target block. Specifically, the predicted signal generator 103 derives reconstructed pixel signals in the same frame from the frame memory 104 and extrapolates these signals to generate the intra-frame predicted signal. The information indicating the method of extrapolation is fed

via line L112 to the entropy encoder 111 to be encoded thereby and then the encoded data is output from the output terminal 112. The intra-frame predicted signal generated in this manner is fed to the subtracter 105. The method of generating the intra-frame predicted signal in the predicted signal generator 103 is the same as the method of H.264 being the conventional technology. The predicted signal with the smallest error is selected from the inter-frame predicted signal and the intra-frame predicted signal obtained as described above, and the selected predicted signal is fed to the subtracter 105.

[0034] Since there are no pictures prior to the first picture, all target blocks thereof are processed by intra-frame prediction. For switching of TV channels, target blocks regularly defined as random access points are processed by intra-frame prediction. These pictures are called intra frames and are also called IDR pictures in H.264.

[0035] The subtracter 105 subtracts the predicted signal (fed via line L103) from the signal of the target block (fed via line L102) to generate a residual signal. This residual signal is transformed by a discrete cosine transform by the transformer 106 to obtain transform coefficients, which are quantized by the quantizer 107. Finally, the entropy encoder 111 encodes the quantized transform coefficients and the encoded data is output along with the information about the prediction method from the output terminal 112.

[0036] For the intra-frame prediction or the inter-frame prediction of the subsequent target block, the signal of the target block, which is compressed is subjected to inverse processing to be reconstructed. Namely, the quantized transform coefficients are inversely quantized

by the de-quantizer 108 and then transformed by an inverse discrete cosine transform by the inverse-transformer 109, to reconstruct a residual signal. The adder 110 adds the reconstructed residual signal to the predicted signal fed via line L103 to reproduce a signal of the target block and the reconstructed signal is stored in the frame memory 104. The present embodiment employs the transformer 106 and the inverse-transformer 109, but it is also possible to use other transform processing instead of these transformers. Depending upon situations, the transformer 106 and the inverse-transformer 109 may be omitted.

[0037] The frame memory 104 is a finite storage and it is impossible to store all reconstructed pictures. Only reconstructed pictures to be used in encoding of the subsequent picture are stored in the frame memory 104. A unit to control this frame memory 104 is the frame memory manager 114. The frame memory manager 114 controls the frame memory 104 via line L115 so as to delete an unnecessary picture (e.g., the oldest picture) out of N reconstructed pictures in the frame memory 104 (where N is 4 in an embodiment, but N may be any predetermined integer) and thereby allow the latest reconstructed picture as a reference picture to be stored. The frame memory manager 114 also receives output order information of each picture and a type of encoding of each picture (intra-frame predictive encoding, inter-frame predictive encoding, or bidirectional predictive encoding) from the input terminal 113, and the reference index via line L112, and the frame memory manager 114 operates based on these pieces of information.

[0038] At the same time, the output order information of each picture

and information of an NAL unit type described below are fed via line L114 to the entropy encoder 111 according to need, in order to be encoded thereby, and the encoded data is output along with the compressed picture data. The output order information is attendant on each picture and may be information indicative of an order of the picture or a time of output of the picture, or an output reference time (temporal reference) of the picture. In the present embodiment, the value of the output order information is directly converted into a binary code. The operation of the frame memory manager 114 in the present embodiment will be described later.

[0039] [About Video Predictive Decoding Device]

Next, a video predictive decoding device according to the present invention will be described. Fig. 2 is a function block diagram showing a configuration of a video predictive decoding device 200 according to an embodiment of the present invention. As shown in Fig. 2, the video predictive decoding device 200 is provided with an input terminal 201, a data analyzer 202, a de-quantizer 203, an inverse-transformer 204, an adder 205, a predicted signal generator 208, a frame memory 207, an output terminal 206, a frame memory manager 209, a controller 210, and a switch 211 as a functional configuration. The operations of the respective function blocks will be described in the operation of the video predictive decoding device 200 below. The de-quantizer 203 and the inverse-transformer 204 correspond to the decoding means. The means associated with decoding is not limited solely to the de-quantizer 203 and the inverse-transformer 204, but may be any other means. Furthermore, the means associated with

decoding may be configured with the de-quantizer 203 only, excluding the inverse-transformer 204.

[0040] The operation of the video predictive decoding device 200 will be described below. Compressed data resulting from compression encoding by the aforementioned method by the video predictive encoding device 100 is input through the input terminal 201. This compressed data contains the residual signal resulting from predictive encoding of each target block obtained by division of a picture into a plurality of blocks, and the information related to the generation of the predicted signal. The information related to the generation of the predicted signal includes the information about block division (size of block), the motion information, the aforementioned reference index, and the information about NAL unit type in the case of the inter-frame prediction, or the information about the extrapolation method from reconstructed surrounding pixels in the case of the intra-frame prediction.

[0041] The data analyzer 202 extracts the residual signal of the target block, the information related to the generation of the predicted signal, the quantization parameter, and the output order information of the picture from the compressed data. The residual signal of the target block is inversely quantized on the basis of the quantization parameter (fed via lines L202 and L211) by the de-quantizer 203. The result is transformed by an inverse discrete cosine transform by the inverse-transformer 204.

[0042] Next, the information related to the generation of the predicted signal is fed via line L206b to the predicted signal generator 208. The

predicted signal generator 208 accesses the frame memory 207, based on the information related to the generation of the predicted signal, to derive a reference signal from a plurality of reference pictures (via line L207) and generate a predicted signal. The predicted signal is fed via line L208 to the adder 205, the adder 205 adds this predicted signal to the reconstructed residual signal to reproduce a target block signal, and the target block signal is output via line L205 from the output terminal 206 and simultaneously stored in the frame memory 207.

[0043] Reconstructed pictures to be used for decoding and reproduction of the subsequent picture are stored in the frame memory 207. The frame memory manager 209 controls the frame memory 207 via line L209a. The frame memory 207 is controlled so that an unnecessary picture (e.g., the oldest picture) is deleted out of N reconstructed pictures stored (where N is 4 in an embodiment, but N may be any predetermined integer) to allow the latest reconstructed picture as a reference picture to be stored.

[0044] The controller 210 operates based on the output order information of the target picture and the information about the encoding type and the NAL unit type of the picture, which are fed to the controller 210 via line L206a. In another situation, the controller 210 can operate based on the reference index fed via line L206a and the information of the frames fed via line L209b and stored in the frame memory. The operation of the controller 210 according to the present invention will be described later.

[0045] The switch 211 is controlled via line L210 by the controller 210 and operates so as to skip decoding of specific frames depending upon

conditions. The operation of the switch 211 according to the present invention will be described later.

[0046] Fig. 3 shows syntax elements 500 of a bit stream. The syntax elements 500 of the bit stream consist of a plurality of syntax elements necessary for decoding of each picture (510, 520, etc.). In a syntax of a picture, attention is focused on three elements below.

1) Network adaptation layer unit type (NUT) or NAL unit type (530)

2) Picture output count (POC) (540)

3) Reference picture set (RPS) (550)

[0047] 1) NUT includes information about a picture type. It should be noted that the present invention can employ other means for signaling a picture type. In the present embodiment, each picture is labeled as one of three kinds of NAL unit types. The NAL unit types are RAS, CRA, and non-RAS as further described below.

[0048] A picture labeled as a RAS (random access skip) picture is skipped so as not to be output, when decoding is started from a CRA picture associated with the RAS picture. On the other hand, when the foregoing CRA picture is not the first picture of a bit stream (or when decoding is not started from the foregoing CRA picture), the video predictive decoding device 200 regards the RAS picture as a non-RAS picture and is configured to decode and output RAS picture in accordance with an output command of the picture.

[0049] A picture labeled as a CRA (clean random access) picture indicates that when decoding of a bit stream is started from the CRA picture associated with the CRA picture, any picture except for the RAS picture, can be decoded without error.

[0050] A picture labeled as a non-RAS picture is assumed to be decoded by the video predictive decoding device 200 and output in accordance with a picture output command. Each CRA picture is assumed to be a non-RAS picture unless otherwise stated.

5 [0051] 2) POC includes information of an order of an output picture.

[0052] 3) RPS includes information of reference pictures used for inter-frame prediction of a current picture. Any reference picture in the decoded picture buffer (DPB) not existing in RPS cannot be used as a reference picture for predictive decoding by a current picture or by
10 any picture.

[0053] The present embodiment has the following features about RPS, in order to ensure that when decoding of a bit stream is started from a CRA picture, every non-RAS picture is correctly decoded.

15 Feature 1: concerning an RPS used by a leading picture, when one or more reference pictures (or at least one reference picture) are RAS pictures or when they are outputted after a CRA picture associated with the picture, the leading picture shall be deemed a RAS picture.

20 Feature 2: every reference picture in an RPS used by a non-RAS picture shall be deemed as a reference picture of a non-RAS picture and a reference picture decoded after a CRA picture associated with the picture.

[0054] Since in the present embodiment each normal picture is handled as a non-RAS picture, any picture not satisfying Features 1 and 2 is not
25 allowed in a bit stream. However, the present invention is not limited only to the leading picture described in Feature 1, but can be equally

applied to every picture. Concerning Feature 2, the present invention can also be applied to a situation where the reference pictures are limited to leading pictures only.

[0055] [Characteristic Operation in Video Predictive Encoding Device 100]

The operation of the video predictive encoding device 100 for generation of a bit stream with the aforementioned features being the point of the present invention will be described using Fig. 4. The video predictive encoding device 100 puts CRA pictures in a fixed period in the bit stream, for implementation of random access. All pictures following one input CRA picture in encoding order are associated with the input CRA picture and encoded according to the steps below, before the next CRA picture is put in.

[0056] It is determined in step 620 whether one or more of reference pictures in the RPS of the picture (i.e. target picture for encoding) are RAS pictures. When one or more of the reference pictures in the RPS of the target picture are RAS pictures (YES), the flow goes to step 650; if not (NO) the flow goes to step 630.

[0057] It is determined in step 630 whether one or more of the reference pictures in the RPS of the target picture are outputted before a CRA picture associated with the target picture. When one or more of the reference pictures in the RPS of the target picture are outputted before the CRA picture associated with the target picture in encoding order (YES), the flow goes to step 650; if not (NO) the flow goes to step 640.

[0058] In step 650, the POC of the target picture is compared with the

POC of the CRA picture associated with the target picture, whereby it is checked whether the target picture is a leading picture. When the POC of the target picture is smaller than the POC of the CRA picture associated with the target picture, the target picture is determined to be a leading picture (YES) and then the flow goes to step 670. Otherwise, the target picture is determined not to be a leading picture (NO); however, the determinations in step 620 and step 630 should be (YES) for only leading pictures, and the determination result that the target picture is not a leading picture (NO) is abnormal; therefore, the flow goes to step 660 to output an error message and then goes to step 680. After the output of the error message in step 660, the processing of Fig. 4 may be terminated as an abnormal end.

[0059] In step 670, the target picture is encoded as a RAS picture and information indicating that the target picture is a RAS picture (NAL unit type: RAS) is encoded. Thereafter, the flow goes to step 680.

[0060] In step 640, the target picture is encoded as a non-RAS picture and information indicating that the target picture is a non-RAS picture (NAL unit type: non-RAS) is encoded. Thereafter, the flow goes to step 680. It is noted herein that the CRA pictures are included in non-RAS pictures unless otherwise stated.

[0061] In steps 640 and 670, the information indicating that the target picture is a RAS picture or a non-RAS picture does not always have to be encoded, but, instead of encoding of the foregoing information, whether the target picture is a RAS picture or a non-RAS picture may be determined by comparison between the reference picture list of each picture and pictures stored in the frame memory 104.

[0062] In step 680 the video predictive encoding device 100 determines whether there is a further picture to be encoded; if there is (YES) the flow returns to step 620 to repeat the processing; if not (NO), the processing of Fig. 4 is terminated.

5 [0063] The sequential processing described above corresponds to the processing of the entire video predictive encoding device 100 in Fig. 1, and among others, the determination processes in steps 620, 630, and 650 are performed by the frame memory manager 114.

[0064] [Characteristic Operation in Video Predictive Decoding Device
10 200]

The video predictive decoding device 200 of the present embodiment operates differently when a decoding process is started from a CRA picture as the first picture of a bit stream, from when the first picture of the bit stream is not a CRA picture. This decoding
15 process returns to a normal decoding process upon decoding of the next CRA picture.

[0065] The operation of the video predictive decoding device 200 for decoding of a bit stream with the aforementioned features of the present invention will be described using Fig. 5.

20 [0066] In step 710, the video predictive decoding device 200 determines, based on the NAL unit type, whether the first picture of the bit stream (i.e., the first picture at a start of decoding of the bit stream) is a CRA picture. When the first picture is not a CRA picture (NO), the flow goes to step 780 where the video predictive decoding device
25 200 decodes each picture according to the normal operation. Namely, in this step 780 a RAS picture is regarded as a non-RAS picture and is

decoded and output according to a command in the picture according to the normal operation. On the other hand, when the first picture of the bit stream is a CRA picture in step 710 (YES), the flow goes to step 720.

5 [0067] The processing from step 720 to step 770 is repeatedly executed for all pictures, during a period immediately before a start of decoding of the next CRA picture, and thereafter, the processing returns to the normal decoding process in step 780. The processing from step 720 to step 770 will be described below.

10 [0068] In step 720, the video predictive decoding device 200 determines whether the picture (i.e. target picture for decoding) is correctly decoded, at a start of decoding of the target picture. Since the bit stream in the present embodiment has Features 1 and 2 described above, the video predictive decoding device 200 can
15 determine whether the target picture can be correctly decoded, using at least one of two methods below. The first method is a method of checking a label of the NAL unit type of the target picture. If the target picture is labeled as a RAS picture, the video predictive decoding device 200 can determine that the target picture cannot be correctly
20 decoded. The second method is a method in which the video predictive decoding device 200 compares the reference pictures in the DPB with the reference picture list of the RPS of the target picture. If any one of the reference pictures in the RPS of the target picture does not exist in the DPB, the video predictive decoding device 200 can
25 determine that the target picture cannot be correctly decoded. When the video predictive decoding device 200 determines that the target

picture can be correctly decoded (YES), using at least one of the first and second methods as described above, the flow goes to step 730; when the device determines that the picture cannot be correctly decoded (NO), the flow goes to step 750.

5 [0069] In step 730, the video predictive decoding device 200 decodes and outputs the target picture in accordance with a command in the target picture. This is also applied to the CRA picture. Thereafter, the flow goes to step 740.

10 [0070] In step 750, the device compares the POC of the target picture with the POC of the CRA picture associated with the target picture, thereby determining whether the target picture is a leading picture. When the POC of the target picture is smaller than the POC of the CRA picture associated with the target picture (YES), the target picture is determined to be a leading picture and the flow goes to step 770
15 described below. Otherwise (NO), the target picture is not a leading picture and can cause an error; therefore, the flow goes to step 760 where the video predictive decoding device 200 outputs an error message and proceeds to step 740. After the output of the error message in step 760, the processing of Fig. 5 may be terminated as an abnormal end. It should be noted as described above that the
20 determination in step 750 is needed only when Feature 1 is limited to leading pictures only.

[0071] In step 770, the video predictive decoding device 200 skips
25 decoding of the target picture by not subjecting the target picture to decoding, and performs a necessary housekeeping process as described below. The necessary housekeeping process herein can be, for

example, a process of labeling the target picture as skipped, with a label indicating that "the picture is unavailable as a reference frame and thus is not output." Thereafter, the flow goes to step 740.

[0072] In step 740, the video predictive decoding device 200 determines whether a picture to be decoded next is a CRA picture, and when the next picture is not a CRA picture (NO), the device returns to step 720 to repeat the processing. On the other hand, when the next picture is a CRA picture (YES), the decoding process according to the present invention (random access decoding process) is no longer necessary after the next CRA picture and therefore the flow goes to step 780 to move into the normal decoding process (process of decoding every picture and outputting it according to output order information).

[0073] The sequential processing described above corresponds to the processing of the entire video predictive decoding device 200 in Fig. 2 and among others the determinations in steps 720 and 750 and the controls in steps 730 and 770 are carried out by the controller 210.

[0074] According to the present embodiment as described above, the video predictive decoding device 200 is able to detect whether a certain picture can be correctly decoded (by use of the label or by comparison with the reference picture set), when decoding is started from the CRA picture at the head of the bit stream. For this reason, the video predictive decoding device 200 can select and discard only a picture that cannot be decoded, instead of discarding all the leading pictures, so as to allow a decodable picture to be used as a reference picture for a subsequent picture, thereby contributing to improvement in prediction

performance.

[0075] In assigning the NAL unit type of RAS to pictures, the video predictive encoding device 100 generates correctly-decodable pictures and undecodable pictures. On the other hand, the video predictive decoding device 200 does not output the undecodable pictures. This makes temporal gaps in between output pictures, which can affect an output rate of frames. The existence of gaps of output is unfavorable for some systems. In the present embodiment, the video predictive encoding device 100 notifies the video predictive decoding device 200 of whether there are gaps associated with the RAS pictures, as additional information by a flag in the CRA picture header or in a video usability syntax (VUI). The video predictive decoding device 200, receiving this flag, can select whether a leading picture with a gap that can be correctly decoded is to be output.

[0076] As another means different from the above, a further restriction may be set on a bit stream so as to avoid a gap at a RAS picture that is outputted after a CRA picture. Namely, the bit stream may be arranged so as to be continuously output without gaps at RAS pictures.

[0077] As still another means, the video predictive decoding device 200 may determine that a leading picture of non-RAS is decoded but not output, independent of the other additional information from the video predictive encoding device 100 or of the output order information of the picture.

[0078] In the present embodiment the labels of the NAL unit types (RAS, CRA, and non-RAS) are detected and used by the video predictive decoding device 200, but the labels of NAL unit types may

be detected and used for execution of processing to discard the RAS picture, when decoding is started from a random access point, in other devices (e.g., a server, appropriate network elements, and so on) in a network. This can save the network bandwidth.

5 [0079] In the present embodiment, each bit stream can include a large number of CRA pictures and there are RAS pictures associated with the respective CRA pictures. When a second CRA picture in decoding order follows a first CRA picture, the RPS of the foregoing second CRA picture is not allowed to include any reference picture decoded
10 before the first CRA picture. This ensures that when the first CRA picture is the first picture of the bit stream, the RAS picture of the second CRA picture is decoded.

[0080] [About Video Predictive Encoding Program and Video Predictive Decoding Program]

15 The invention of the video predictive encoding device 100 can also be interpreted as the invention of a video predictive encoding program for letting a computer function as the video predictive encoding device 100. Likewise, the invention of the video predictive decoding device 200 can also be interpreted as the invention of a video
20 predictive decoding program for letting a computer function as the video predictive decoding device 200.

[0081] The video predictive encoding program and the video predictive decoding program are provided, for example, as stored in a storage medium. Examples of such storage media include flexible disks,
25 CD-ROMs, USB memories, DVDs, semiconductor memories, and so on.

[0082] Fig. 8 shows modules of the video predictive encoding program for letting a computer function as the video predictive encoding device 100. As shown in Fig. 8, the video predictive encoding program P100 is provided with an input module P101, an encoding module P102, a reconstruction module P103, a picture storage module P104, and a control module P105.

[0083] Fig. 9 shows modules of the video predictive decoding program for letting a computer function within the video predictive decoding device 200. As shown in Fig. 9, the video predictive decoding program P200 is provided with an input module P201, a reconstruction module P202, a picture storage module P203, and a control module P204.

[0084] The video predictive encoding program P100 and the video predictive decoding program P200 configured as described above can be stored in a storage medium 10 shown in Figs. 6 and 7 and are executed by a computer 30 described below.

[0085] Fig. 6 is a drawing showing a hardware configuration of a computer for executing a program stored in a storage medium and Fig. 7 a general view of a computer for executing a program stored in a storage medium. The computer embraces a DVD player, a set-top box, a cell phone, etc. provided with a CPU and configured to perform processing and control by software.

[0086] As shown in Fig. 6, the computer 30 is provided with a reading device 12 such as a flexible disk drive unit, a CD-ROM drive unit, or a DVD drive unit, a working memory (RAM) 14 on which an operating system is resident, a memory 16 for storing programs stored in the

storage medium 10, a monitor unit 18 like a display, a mouse 20 and a keyboard 22 as input devices, a communication device 24 for transmission and reception of data or the like, and a CPU 26 for controlling execution of programs. When the storage medium 10 is put into the reading device 12, the computer 30 becomes accessible to the video predictive encoding program stored in the storage medium 10, through the reading device 12 and becomes able to operate as the video predictive encoding device according to the present invention, through execution of the video predictive encoding program. Similarly, when the storage medium 10 is put into the reading device 12, the computer 30 becomes accessible to the video predictive decoding program stored in the storage medium 10, through the reading device 12 and becomes able to operate as the video predictive decoding device according to the present invention, through execution of the video predictive decoding program.

[0087] As shown in Fig. 7, the video predictive encoding program or the video predictive decoding program may be one provided in the form of computer data signal 40 superimposed on a carrier wave, through a network. In this case, the computer 30 can execute the video predictive encoding program or the video predictive decoding program after the video predictive encoding program or the video predictive decoding program received by the communication device 24 is stored into the memory 16.

List of Reference Signs

[0088] 10: storage medium; 30: computer; 100: video predictive encoding device; 101: input terminal; 102: block divider; 103:

predicted signal generator; 104: frame memory; 105: subtracter; 106:
transformer; 107: quantizer; 108: de-quantizer; 109:
inverse-transformer; 110: adder; 111: entropy encoder; 112: output
terminal; 113: input terminal; 114: frame memory manager; 200: video
5 predictive decoding device; 201: input terminal; 202: data analyzer;
203: de-quantizer; 204: inverse-transformer; 205: adder; 206: output
terminal; 207: frame memory; 208: predicted signal generator; 209:
frame memory manager; 210: controller; P100: video predictive
encoding program; P101: input module; P102: encoding module; P103:
10 reconstruction module; P104: picture storage module; P105: control
module; P200: video predictive decoding program; P201: input
module; P202: reconstruction module; P203: picture storage module;
P204: control module.

CLAIMS

1. A video predictive decoding device comprising:

input means that inputs a bitstream including compressed picture data for a plurality of pictures constituting a video sequence, where each picture has a network abstraction layer unit type that identifies said picture as one of a plurality of picture types including random access picture, random access skipped (RAS) leading picture and non-RAS leading picture;

reconstruction means which decodes the compressed picture data to reconstruct pictures based on the picture types; and

output means which outputs the reconstructed pictures;

wherein

1) a random access picture is the first picture in the bitstream in decoding order when the decoding process, which starts at any random access picture in the bitstream, is started from said random access picture;

2) RAS leading picture is the picture which precedes the associated random access picture in output order, and is not decodable when the associated random access picture is the first picture in the bitstream in decoding order;

3) non-RAS leading picture is the picture which precedes the associated random access picture in output order, and is decodable,

and wherein

a reference picture set of the non-RAS leading picture including reference pictures used for inter prediction of the non-RAS leading picture does not include any of a RAS leading picture or a picture that precedes the associated random access picture in decoding order, and

a reference picture set of a second random access picture does not include any picture preceding a first random access picture in decoding order when the second random access picture is decoded after the first random access picture.

2. A video predictive decoding device according to claim 1, wherein the reference picture set of the second random access picture does not include any picture preceding the first random access picture in decoding order when the second random access picture is decoded after the first random access picture, the reference picture set of the second random access picture including reference pictures used for inter prediction of an other picture that succeeds the second random access picture in decoding order.

3. A video predictive decoding method executed by a video predictive decoding device, comprising:

an input step of inputting a bitstream including compressed picture data for a plurality of pictures constituting a video sequence, where each picture has a network abstraction layer unit type that identifies said picture as one of a plurality of picture types including random access picture, random access skipped (RAS) leading picture and non-RAS leading picture;

a reconstruction step of decoding the compressed picture data to reconstruct pictures based on the picture types; and

an output step of outputting the reconstructed pictures;

wherein

1) a random access picture is the first picture in the bitstream in decoding order when the decoding process, which starts at any random access picture in the bitstream, is started from said random access picture;

2) RAS leading picture is the picture which precedes the associated random access picture in output order, and is not decodable when the associated random access picture is the first picture in the bitstream in decoding order;

3) non-RAS leading picture is the picture which precedes the associated random access picture in output order, and is decodable,

and wherein

a reference picture set of the non-RAS leading picture including reference pictures used for inter prediction of the non-RAS leading picture does not include any of a RAS leading picture or a picture that precedes the associated random access picture in decoding order, and

a reference picture set of a second random access picture does not include any picture preceding a first random access picture in decoding order when the second random access picture is decoded after the first random access picture.

4. A video predictive decoding method according to claim 3, wherein the reference picture set of the second random access picture does not include any picture preceding the first random access picture in decoding order when the second random access picture is decoded after the first random access picture, the reference picture set of the second random access picture including reference pictures used for inter prediction of an other picture that succeeds the second random access picture in decoding order.

5. A video predictive decoding program for letting a computer function as:

input means that inputs a bitstream including compressed picture data for a plurality of pictures constituting a video sequence, where each picture has a network abstraction layer unit type that identifies said picture as one of a plurality of picture types including random access picture, random access skipped (RAS) leading picture and non-RAS leading picture;

reconstruction means which decodes the compressed picture data to reconstruct pictures based on the picture types; and

output means which outputs the reconstructed pictures;

wherein

1) a random access picture is the first picture in the bitstream in decoding order when the decoding process, which starts at any random access picture in the bitstream, is started from said random access picture;

2) RAS leading picture is the picture which precedes the associated random access picture in output order, and is not decodable when the associated random access picture is the first picture in the bitstream in decoding order;

3) non-RAS leading picture is the picture which precedes the associated random access picture in output order, and is decodable,

and wherein

a reference picture set of the non-RAS leading picture including reference pictures used for inter prediction of the non-RAS leading picture does not include any of a RAS leading picture or a picture that precedes the associated random access picture in decoding order, and

a reference picture set of a second random access picture does not include any picture preceding a first random access picture in decoding order when the second random access picture is decoded after the first random access picture.

6. A video predictive decoding program according to claim 5, wherein the reference picture set of the second random access picture does not include any picture preceding the first random access picture in decoding order when the second random access picture is decoded after the first random access picture, the reference picture set of the second random access picture including reference pictures used for inter prediction of an other picture that succeeds the second random access picture in decoding order.