(12) STANDARD PATENT (11) Application No. **AU 2020201376 B2**
(19) AUSTRALIAN PATENT OFFICE

(54) Title
Video prediction encoding device, video prediction encoding method, video prediction encoding program, video prediction decoding device, video prediction decoding method, and video prediction decoding program

(51) International Patent Classification(s)
H04N 19/50 (2014.01)

(21) Application No: **2020201376** (22) Date of Filing: **2020.02.25**

(43) Publication Date: **2020.03.12**
(43) Publication Journal Date: **2020.03.12**
(44) Accepted Journal Date: **2021.06.10**

(62) Divisional of:
2018208737

(71) Applicant(s)
NTT DOCOMO, INC.

(72) Inventor(s)
FUJIBAYASHI, Akira;BOON, Choong Seng;TAN, Thiow Keng

(74) Agent / Attorney
Spruson & Ferguson, GPO Box 3898, Sydney, NSW, 2001, AU

(56) Related Art
US 2005/0025250 A1

VIDEO PREDICTION ENCODING DEVICE, VIDEO PREDICTION ENCODING METHOD, VIDEO PREDICTION ENCODING PROGRAM, VIDEO PREDICTION DECODING DEVICE, VIDEO PREDICTION DECODING METHOD, AND VIDEO PREDICTION DECODING PROGRAM

ABSTRACT

A video predictive encoding device comprising: a CPU programmable to implement video predictive encoding; a receiver executable with the programmed CPU to receive a plurality of pictures forming a video sequence; an encoder executable with the programmed CPU to encode the received pictures by either intra-frame prediction or inter-frame prediction to generate compressed picture data; a decoder executable with the programmed CPU to decode the compressed picture data to reconstruct the pictures; a frame memory executable with the programmed CPU to store one or more of the reconstructed pictures as reference pictures to be used for encoding a subsequent picture; and a memory manager executable with the programmed CPU to manage the frame memory, the memory manager being executable to retrieve (i) a target picture frame size indicative of a frame size of a target picture, (ii) a maximum frame size (MaxLumaFS) defined in advance by level information and indicative of a maximum size of a reconstructed picture storable in the frame memory, and (iii) a size of the frame memory (MFSBuffer) expressed by a preset maximum number of reconstructed pictures storable in the frame memory, wherein the memory manager is further executable to: compare a halved MaxLumaFS with the target picture frame size; upon a determination that the halved MaxLumaFS is equal to or larger than the target picture frame size, double the MFSBuffer; and set an adaptive maximum number (maxDPBsize) of the reconstructed pictures storable in the frame memory equal to a doubled MFSBuffer, and wherein a number of the reconstructed pictures stored in the frame memory is no more than a number set in the maxDPBsize.

VIDEO PREDICTION ENCODING DEVICE, VIDEO PREDICTION ENCODING METHOD, VIDEO PREDICTION ENCODING PROGRAM, VIDEO PREDICTION DECODING DEVICE, VIDEO PREDICTION DECODING METHOD, AND VIDEO PREDICTION DECODING PROGRAM

[0001] The present application is a divisional application of Australian Patent Application No. 2018208737, filed 27 July 2018, which in turn is a divisional of Australian Patent Application No. 2016210706, filed 4 August 2016, which itself is a divisional of Australian Patent Application No. 2013284789, filed 24 April 2013. The contents of Australian Patent Application No. 2018208737, Australian Patent Application No. 2016210706 and Australian Patent Application No. 2013284789 is incorporated herein by reference in their entirety.

TECHNICAL FIELD

[0002] The present invention relates to video predictive encoding and decoding devices, methods, and programs and, more particularly, to devices, methods, and programs to manage the maximum number of reconstructed pictures storable in a decoded picture buffer.

BACKGROUND ART

[0003] The video compression technologies are used for efficiently transmitting and storing video data. MPEG1-4 and H.261 to H.264 are widely used video data compression standards.

[0004] In these video compression standards, a picture to be encoded is divided into a plurality of blocks, which are encoded and then decoded. In order to increase the coding efficiency, below described prediction coding is used. In intra-frame prediction, a predictive signal is generated using a signal of a reconstructed neighboring picture (a signal restored from a previously compressed picture data) which is present in the frame including the target block. By subtracting the predictive signal from a signal of the target block, a difference between them is obtained and encoded. In inter-frame prediction, a reconstructed picture signal, which is

202020137 6 14 May 2021

present in a frame different from the frame including a target block, is searched for a displacement of the signal. A predictive signal is generated to compensate the displacement. By subtracting the predictive signal from the signal of the target block, a difference between them is obtained and encoded. The reconstructed picture which is the subject of the motion search and compensation is referred to as a reference picture.

[0005] In bidirectional inter-frame prediction, not only is a past picture referenced, but also a future picture is referenced which is ordered for display after the target picture (the future pictures need to be encoded and reconstructed prior to encoding of the target picture). Then, predictive signals acquired from the past picture and the future picture are averaged. This prediction method is effective to predict an object not present in the past but thrown in a future frame and to reduce noise included in the two predictive signals.

[0006] Furthermore, in the inter-frame prediction defined in H.264, a plurality of reference pictures which have been encoded and then reconstructed are subjected to motion search, and the predictive signal with the smallest error is selected as an optimum predictive signal for the target block. A difference is calculated between the pixel signal of the target block and the optimum predictive signal and then subjected to a discrete cosine transform, quantization, and entropy encoding. At the same time, information is encoded which identifies the selected reference picture and the region in the selected reference picture from which the optimum predictive signal for the target block is acquired (referred to as a "reference index" and a "motion vector," respectively).

[0007] In H.264, a plurality of reconstructed pictures may be referenced. These reconstructed pictures are stored, as reference pictures to be used in prediction, in a decoded picture buffer (DPB), which is a picture buffer memory. The size of the decoded picture buffer (DPB) is defined by a profile and a level, and defined as a bit count, instead of the number of reference pictures. Even with the same profile and level, the number of the storable reference pictures varies according to the frame size of pictures. For example, in the case where the profile is main (Main) and the level is 3.2, the maximum size of the picture buffer (MaxDPBSize:

2020201376 14 May 2021

Maximum Decoded Picture Buffer Size) for storage of reference pictures used for prediction is defined as 7680.0×1024 [bytes]. Therefore, the number of reconstructed pictures storable in the decoded picture buffer (DPB) is 5 if the pictures are of 1280×720 and 4:2:0, and the maximum number of reconstructed pictures storable in the decoded picture buffer (DPB) is 4 if the pictures are of 1280×1024 and 4:2:0. Figs. 1(a), 1(b) and 1(c) show pictures arranged in the decoded picture buffer in which the frame sizes of the pictures determine the maximum number of storable reconstructed pictures, which is 4 (Fig. 1(a)), 5 (Fig. 1(b)), or 6 (Fig. 1(c)). Memory pointers are provided adaptively to the frame size of the reconstructed pictures in the picture buffer memory prepared in advance, whereby an adaptive memory arrangement is achieved in the picture buffer memory.

[0008] Non Patent Literature 1: "H.264: Advanced video coding for generic audiovisual services," Joint Video Team of ITU-T VCEG and ISO/IEC MPEG, ITU-T Rec. H.264 and ISO/IEC14496-10 (MPEG4-Part 10), November 2007.

[0009] Since the decoding device which complies with the foregoing requirements of H.264 needs to be able to decode pictures of all frame sizes defined by the profiles and levels which fall within the range supported by the decoding device, it is not possible to arrange memories in the picture buffer memory and store reconstructed pictures in these memories in an fixed manner. For this reason, it is necessary to vary, according to the frame size of pictures to be decoded, the memory pointers pointing the positions of memories for storing the reconstructed pictures, resulting in making the memory control and implementation complicated.

[0010] A makeshift solution to this problem is adopting a technique in which the maximum number of reconstructed pictures storable in the picture buffer memory (max_dec_pic_buffering) is fixed, regardless of the frame sizes of decoded pictures. Under this solution, since the memory arrangement of the decoded picture buffer (DPB) is fixedly determined by the maximum frame size as shown in Fig. 2(a), the memory pointers can also be fixed. Namely, the memory control becomes easier compared to the arrangement in which the memory pointers need to be variably controlled. However, when the frame size of decoded

pictures is smaller than the maximum frame size or when the vertical width of pictures is half of that of the frame pictures as in the case of interlaced field pictures, implementation of the above solution could generate unused memory regions as shown in Fig. 2(b) and prevent the memory from being effectively used. Furthermore, the solution could deprive a room for improving the coding efficiency which could be realized by storing more reconstructed pictures to increase the number of reference pictures.

SUMMARY

[0011] It is the object of the present invention to substantially overcome or at least ameliorate one or more of the above disadvantages.

[0012] Aspects of the present disclosure provide encoding and decoding methods, devices, and programs in which the maximum number of reconstructed pictures storable in the picture buffer memory is determined, according to the frame size of reconstructed pictures, thereby enabling the efficient use of the picture buffer memory and further improvement of the coding efficiency.

[0013] In one aspect, the present invention provides a video predictive encoding device comprising: a CPU programmable to implement video predictive encoding; a receiver executable with the programmed CPU to receive a plurality of pictures forming a video sequence; an encoder executable with the programmed CPU to encode the received pictures by either intra-frame prediction or inter-frame prediction to generate compressed picture data; a decoder executable with the programmed CPU to decode the compressed picture data to reconstruct the pictures; a frame memory executable with the programmed CPU to store one or more of the reconstructed pictures as reference pictures to be used for encoding a subsequent picture; and a memory manager executable with the programmed CPU to manage the frame memory, the memory manager being executable to retrieve (i) a target picture frame size indicative of a frame size of a target picture, (ii) a maximum frame size (MaxLumaFS) defined in advance by level information and indicative of a maximum size of a reconstructed picture storable in the frame memory, and (iii) a size of the frame memory (MFSBuffer) expressed by a

2020201376 14 May 2021

preset maximum number of reconstructed pictures storable in the frame memory, wherein the memory manager is further executable to: compare a halved MaxLumaFS with the target picture frame size; upon a determination that the halved MaxLumaFS is equal to or larger than the target picture frame size, double the MFSBuffer; and set an adaptive maximum number (maxDPBsize) of the reconstructed pictures storable in the frame memory equal to a doubled MFSBuffer, and wherein a number of the reconstructed pictures stored in the frame memory is no more than a number set in the maxDPBsize.

[0014] In one aspect, the present invention provides a video predictive decoding device comprising: a CPU programmed to implement video predictive decoding; a receiver executable with the CPU to receive encoded data comprising compressed picture data representative of a compressed form of a plurality of pictures forming a video sequence, wherein the plurality of pictures are encoded by either intra-frame prediction or inter-frame prediction; a decoder executable with the CPU to decode the compressed picture data to reconstruct the pictures; a decoded picture buffer (DPB) executable with the CPU to store one or more of the reconstructed pictures as reference pictures to be used for decoding a subsequent picture; and a buffer manager executable with the programmed CPU to manage the DPB, the buffer manager being executable to retrieve (i) a target picture frame size indicative of a frame size of a target picture, (ii) a maximum frame size (MaxLumaFS) defined in advance by level information and indicative of a maximum size of a reconstructed picture storable in the DPB, and (iii) a size of the frame memory (MFSBuffer) expressed by a preset maximum number of reconstructed pictures storable in the DPB, wherein the buffer manager is executable to: compare a halved MaxLumaFS with the target picture frame size; upon a determination that the halved MaxLumaFS is equal to or larger than the target picture frame size, double the MFSBuffer; and set an adaptive maximum number (maxDPBsize) of reconstructed pictures storable in the DPB equal to a doubled MFSBuffer, and wherein a number of the reconstructed pictures stored in the DPB is no more than a number set in the maxDPBsize.

[0014a] In one aspect, the present invention provides a video predictive encoding method executed by a video predictive encoding device, comprising: receiving a plurality of pictures forming a video sequence; encoding the received pictures by either intra-frame prediction or inter-frame prediction to generate compressed picture data; decoding the compressed picture data to reconstruct the pictures; storing one or more of the reconstructed pictures in a frame memory as reference pictures to be used for encoding a subsequent picture; retrieving (i) a target picture frame size indicative of a frame size of a target picture, (ii) a maximum frame size (MaxLumaFS) defined in advance by level information and indicative of a maximum size of a reconstructed picture storable in the frame memory, and (iii) a size of the frame memory (MFSBuffer) expressed by a preset maximum number of reconstructed pictures storable in the frame memory; comparing a halved MaxLumaFS with the target picture frame size; upon a determination that the halved MaxLumaFS is equal to or larger than the target picture frame size, doubling the MFSBuffer; and setting an adaptive maximum number (maxDPBsize) of the reconstructed pictures storable in the frame memory equal to a doubled MFSBuffer, wherein a number of the reconstructed pictures stored in the frame memory is no more than a number set in the maxDPBsize.

[0014b] In one aspect, the present invention provides a video predictive decoding method executed by a video predictive decoding device, comprising: receiving encoded data comprising compressed picture data representative of a compressed form of a plurality of pictures forming a video sequence, wherein the plurality of pictures are encoded by either intra-frame prediction or inter-frame prediction; decoding the compressed form of the plurality of pictures to reconstruct the plurality of pictures; storing in a decoded picture buffer (DPB) one or more of the reconstructed pictures as reference pictures to be used for decoding a subsequent picture; retrieving (i) a target picture frame size indicative of a frame size of a target picture, (ii) a maximum frame size (MaxLumaFS) defined in advance by level information and indicative of a maximum size of a reconstructed picture storable in the DPB, and (iii) a size of the frame memory (MFSBuffer) expressed by a preset maximum number of reconstructed pictures

2020201376 14 May 2021

storable in the DPB; comparing a halved MaxLumaFS with the target picture frame size; upon a determination that the halved MaxLumaFS is equal to or larger than the target picture frame size, doubling the MFSBuffer; and setting an adaptive maximum number (maxDPBsize) of the reconstructed pictures storable in the DPB equal to a doubled MFSBuffer, wherein a number of the reconstructed pictures stored in the DPB is no more than a number set in the maxDPBsize.

[0014c] In one aspect, the present invention provides a non-transitory storage medium that stores a video predictive decoding program which programs a CPU to execute: a receiver operable to receive encoded data comprising compressed picture data representative of a compressed form of a plurality of pictures forming a video sequence, wherein the plurality of pictures are encoded by either intra-frame prediction or inter-frame prediction; a decoder operable to decode the compressed picture data to reconstruct the pictures; a decoded picture buffer (DPB) operable to store one or more of the reconstructed pictures as reference pictures to be used for decoding a subsequent picture; and a buffer manager operable to manage the DPB, the buffer manager being executable to retrieve (i) a target picture frame size indicative of a frame size of a target picture, (ii) a maximum frame size (MaxLumaFS) defined in advance by level information and indicative of a maximum size of a reconstructed picture storable in the DPB, and (iii) a size of the frame memory (MFSBuffer) expressed by a preset maximum number of reconstructed pictures storable in the DPB; and wherein the buffer manager further operable to: compare a halved MaxLumaFS with the target picture frame size; upon a determination that the halved MaxLumaFS is equal to or larger than the target picture frame size, double the MFSBuffer; and set an adaptive maximum number (maxDPBsize) of the reconstructed pictures storable in the DPB equal to a doubled MFSBuffer, and wherein a number of the reconstructed pictures stored in the DPB is no more than a number set in the maxDPBsize.

[0015] An aspect of the present disclosure provides a video predictive encoding method executed by a video predictive encoding device, comprising: receiving a plurality of pictures forming a video sequence; encoding the received pictures by either intra-frame prediction or inter-frame prediction to generate compressed picture data; decoding the compressed picture

2020201376 14 May 2021

data to reconstruct the pictures; storing one or more of the reconstructed pictures in a frame memory as reference pictures to be used for encoding a subsequent picture; retrieving (i) a target picture frame size indicative of a frame size of a target picture, (ii) a maximum frame size (MaxLumaFS) defined in advance by level information and indicative of a maximum size of a reconstructed picture storable in the frame memory, and (iii) a size of the frame memory (MFSBuffer) expressed by a preset maximum number of reconstructed pictures storable in the frame memory; comparing a halved MaxLumaFS with the target picture frame size; upon a determination that the halved MaxLumaFS is equal to or larger than the target picture frame size, doubling the MFSBuffer; and setting an adaptive maximum number (maxDPBsize) of the reconstructed pictures storable in the frame memory equal to a doubled MFSBuffer, wherein a number of the reconstructed pictures stored in the frame memory is no more than a number set in the maxDPBsize.

[0015a] An aspect of the present disclosure provides a non-transitory storage medium that stores a video predictive decoding program which programs a CPU to execute: a receiver operable to receive encoded data comprising compressed picture data representative of a compressed form of a plurality of pictures forming a video sequence, wherein the plurality of pictures are encoded by either intra-frame prediction or inter-frame prediction; a decoder operable to decode the compressed picture data to reconstruct the pictures; a decoded picture buffer (DPB) operable to store one or more of the reconstructed pictures as reference pictures to be used for decoding a subsequent picture; and a buffer manager operable to manage the DPB, the buffer manager being executable to retrieve (i) a target picture frame size indicative of a frame size of a target picture, (ii) a maximum frame size (MaxLumaFS) defined in advance by level information and indicative of a maximum size of a reconstructed picture storable in the DPB, and (iii) a size of the frame memory (MFSBuffer) expressed by a preset maximum number of reconstructed pictures storable in the DPB; and wherein the buffer manager further operable to: compare a halved MaxLumaFS with the target picture frame size; upon a determination that the halved MaxLumaFS is equal to or larger than the target picture frame size, double the MFSBuffer; and set an adaptive maximum number (maxDPBsize) of the

2020201376 14 May 2021

reconstructed pictures storable in the DPB equal to a doubled MFSBuffer, and wherein a number of the reconstructed pictures stored in the DPB is no more than a number set in the maxDPBsize.

[0016] In this video predictive encoding device, the maximum number of reconstructed pictures storable in the picture buffer memory is determined, based on the relationship between the frame size associated with respective pictures forming the video sequence or the compressed picture data and the maximum frame size specified as an encoding constraint. More specifically, for example, when the frame size of the input pictures is larger than 1/(2^L) (where L is an integer of one or more) of the maximum frame size specified as an encoding constraint, the maximum number of storable reconstructed pictures is set to N (where N is an integer of one or more), whereas when the frame size of the input pictures is not more than 1/(2^L) of the maximum frame size, the memory buffer amount available for storing reconstructed pictures is controlled to allow storage of up to the ((2^L)×N) number of reconstructed pictures. In this configuration, when the frame size of reconstructed pictures is not more than 1/(2^L) of the maximum frame size specified as an encoding setting, the maximum number of storable reconstructed pictures can be increased, and thus the number of storable reference pictures can also be increased to thereby improve the coding efficiency.

[0017] In a video predictive encoding device according to another aspect of the present disclosure, the memory control means fixes memory pointers pointing storage positions of reconstructed pictures in the memory when determining the maximum number of reconstructed pictures storable in the frame memory.

[0018] In this video predictive encoding device, the memory pointers pointing the storage positions of the respective reconstructed pictures can be fixed, for example, at positions equal to the maximum frame size and 1/(2^L) thereof. Therefore, control is not needed to vary the memory pointers depending upon the frame sizes, and simpler implementation, more efficient memory control can be realized. Since pointer intervals are limited to multiples of 1/(2^L), the positions of the memory pointers may be calculated with a shift operation, and thus the calculation cost can be reduced.

2020201376 14 May 2021

[0019] A video predictive decoding device according to an aspect of the present disclosure comprises input means that inputs encoded data comprising compressed picture data which includes information on a target picture frame size, a maximum number of storable reconstructed pictures, and a maximum frame size specified as an encoding constraint, wherein a plurality of pictures constituting a video sequence has been compressed by either intra-frame prediction or inter-frame prediction. The video predictive decoding device further comprises decoding means that decodes the compressed picture data and the compressed picture data constituting the information on the target picture frame size, the maximum number of storable reconstructed pictures, and the maximum frame size to reconstruct pictures and the information on the target picture frame size, the maximum number of storable reconstructed pictures, and the maximum frame size. The device further comprises picture storage means that stores one or more of the reconstructed pictures as reference pictures to be used for decoding a subsequent picture and memory control means that controls the picture storage means, wherein when a frame size of the input pictures is not more than 1/(2^L) (where L is an integer of one or more) of the maximum frame size of pictures, the memory control means determines that (2^L) times the maximum number of reconstructed pictures are storable in a frame memory .

[0020] In this video predictive decoding device, the maximum number of reconstructed pictures storable in the picture buffer memory is determined based on the relationship between the frame size associated with respective pictures forming the video sequence or the compressed picture data and the maximum frame size specified as an encoding constraint. More specifically, for example, when the frame size of the input pictures is larger than 1/(2^L) (where L is an integer of one or more) of the maximum frame size specified as an encoding constraint, the maximum number of storable reconstructed pictures is set to N (where N is an integer of one or more), whereas when the frame size of the input pictures is not more than 1/(2^L) of the maximum frame size, it is determined that up to the (2^L) x N number of reconstructed pictures are storable. In this configuration, when the frame size of reconstructed pictures is not more than 1/(2^L) of the maximum frame size specified as an encoding setting, the maximum number of storable reconstructed pictures can be increased, and thus the number of storable reference

pictures can also be increased to thereby improve the coding efficiency.

[0021] In a video predictive decoding device according to another aspect of the present disclosure, the memory control means fixes the memory pointers pointing storage positions of reconstructed pictures in the frame memory, when determining the maximum number of reconstructed pictures storable in the frame memory.

[0022] In this video predictive decoding device, the memory pointers pointing the storage positions of the reconstructed pictures can be fixed, for example, at positions equal to the maximum frame size and 1/(2^L) thereof. Therefore, control is not needed to vary the memory pointers depending upon the frame sizes, and simpler implementation, more efficient memory control can be realized. Since pointer intervals are limited to multiples of 1/(2^L), the positions of the memory pointers can be calculated with a shift operation, and therefore the operation cost can be reduced.

[0023] A video prediction encoding method according to an aspect of the present disclosure is executed by a video predictive encoding device. The method comprises an input step of inputting a plurality of pictures forming a video sequence, and an encoding step of encoding the pictures by either intra-frame prediction or inter-frame prediction to generate compressed picture data including information on a target picture frame size, a maximum number of storable reconstructed pictures, and a maximum frame size specified as an encoding constraint, wherein the compressed picture data is compressed along with encoding specification data regarding the information on the target picture frame size, the maximum number of storable reconstructed pictures, and the maximum frame size. The method further comprises a decoding step of decoding the compressed picture data to reconstruct the pictures, a picture storage step of storing one or more of the reconstructed pictures as reference pictures to be used for encoding a subsequent picture, and a memory control step of controlling a picture buffer memory in the picture storage step, wherein when the frame size of the input pictures is not more than 1/(2^L) (where L is an integer of one or more) of the maximum frame size of pictures, the memory control step determines that (2^L) times the maximum number of reconstructed pictures are

2020201376 14 May 2021

storable in the frame memory.

[0024] A video predictive encoding program according to an aspect of the present disclosure comprises an input module that inputs a plurality of pictures forming a video sequence and an encoding module that encodes the pictures by either intra-frame prediction or inter-frame prediction to generate compressed picture data including information on a target picture frame size, a maximum number of storable reconstructed pictures, and a maximum frame size specified as an encoding constraint and encodes the compressed picture data along with encoding specification data regarding the information on the target picture frame size, the maximum number of storable reconstructed pictures, and the maximum frame size. The program further comprises a decoding module that decodes the compressed picture data to reconstruct pictures, a picture storage module that stores one or more of the reconstructed pictures as reference pictures to be used for encoding a subsequent picture, and a memory control module that controls a picture buffer memory in the picture storage module, wherein when the frame size of the input pictures is not more than 1/(2^L) (where L is an integer of one or more) of the maximum frame size of pictures, the memory control module determines that (2^L) times the maximum number of reconstructed pictures are storable in the frame memory.

[0025] The video predictive encoding method and the video predictive encoding program can achieve the same effects as the aforementioned video predictive encoding device.

[0026] A video predictive decoding method according to an aspect of the present disclosure is executed by a video predictive decoding device. The method comprises an input step of inputting encoded data comprising compressed picture data which includes a target picture frame size, a maximum number of storable reconstructed pictures, and a maximum frame size specified as an encoding constraint, wherein a plurality of pictures constituting a video sequence is compressed by either intra-frame prediction or inter-frame prediction. The method further comprises a decoding step of decoding the compressed picture data and the compressed picture data constituting the information on the target picture frame size, the maximum number of storable reconstructed pictures, and the maximum frame size to reconstruct pictures and the

information on the target picture frame size, the maximum number of storable reconstructed pictures, and the maximum frame size. The method further comprises a picture storage step of storing, in picture storage means, one or more of the reconstructed pictures as reference pictures to be used for decoding a subsequent picture and a memory control step of controlling the picture storage means, wherein when the frame size of the input pictures is not more than 1/(2^L) (where L is an integer of one or more) of the maximum frame size of pictures, the memory control step determines that (2^L) times the maximum number of reconstructed pictures are storable in a frame memory.

[0027] A video predictive decoding program according to an aspect of the present disclosure comprises an input module that inputs encoded data comprising compressed picture data which includes a target picture frame size, a maximum number of storable reconstructed pictures, and a maximum frame size specified as an encoding constraint, wherein a plurality of pictures constituting a video sequence is compressed either intra-frame prediction or inter-frame prediction. The program further comprises a decoding module that decodes the compressed picture data and that compressed picture data constituting the information on the target picture frame size, the maximum number of storable reconstructed pictures, and the maximum frame size to reconstruct pictures and the information on the target picture frame size, the maximum number of storable reconstructed pictures, and the maximum frame size. The program further comprises a picture storage module that stores one or more of the reconstructed pictures as reference pictures to be used for decoding a subsequent picture and a memory control module that controls the picture storage module, wherein when the frame size of the input pictures is not more than 1/(2^L) (where L is an integer of one or more) of the maximum frame size of pictures, the memory control module determines that (2^L) times the maximum number of reconstructed pictures are storable in a frame memory.

[0028] The video prediction decoding method and the video prediction decoding program can achieve the same effects as the aforementioned video prediction decoding device can achieve.

[0029] The video predictive encoding and decoding devices, methods, and programs according

to aspects of the present disclosure perform determining the maximum number of reconstructed pictures storable in the picture buffer memory, according to the frame size of the reconstructed pictures, thereby achieving the efficient use of the picture buffer memory and also improving the encoding efficiency.

BRIEF DESCRIPTION OF THE DRAWINGS

[0029a] Example embodiments should become apparent from the following description, which is given by way of example only, of at least one preferred but non-limiting embodiment, described in connection with the accompanying figures.

[0030] Figs. 1(a) and 1(b) are drawings showing picture buffer memories, the number of storable pictures, and positions of variable memory pointers.

Figs. 2(a) and 2(b) are drawings for explaining a problem of fixed memory pointers in the picture buffer memory.

Fig. 3 is a block diagram showing a video predictive encoding device according to an embodiment of the present invention.

Fig. 4 is a block diagram showing a video predictive decoding device according to an embodiment of the present invention.

Fig. 5 is a flowchart showing first video predictive encoding and decoding methods according to an embodiment of the present invention.

Figs. 6(a) and 6(b) are schematic drawings for explaining processing of the first video predictive encoding and decoding methods according to the embodiment of the present invention.

Fig. 7 is a drawing showing a hardware configuration of a computer for executing a program recorded in a recording medium.

2020201376 25 Feb 2020

CLAIMS

1. A video predictive encoding device comprising:

a CPU programmable to implement video predictive encoding;

a receiver executable with the programmed CPU to receive a plurality of pictures forming a video sequence;

an encoder executable with the programmed CPU to encode the received pictures by either intra-frame prediction or inter-frame prediction to generate compressed picture data;

a decoder executable with the programmed CPU to decode the compressed picture data to reconstruct the pictures;

a frame memory executable with the programmed CPU to store one or more of the reconstructed pictures as reference pictures to be used for encoding a subsequent picture; and

a memory manager executable with the programmed CPU to manage the frame memory, the memory manager being executable to retrieve (i) a target picture frame size indicative of a frame size of a target picture, (ii) a maximum frame size (MaxLumaFS) defined in advance by level information and indicative of a maximum size of a reconstructed picture storable in the frame memory, and (iii) a size of the frame memory (MFSBuffer) expressed by a preset maximum number of reconstructed pictures storable in the frame memory,

wherein the memory manager is further executable to:

compare a halved MaxLumaFS with the target picture frame size;

upon a determination that the halved MaxLumaFS is equal to or larger than the target picture frame size, double the MFSBuffer; and

set an adaptive maximum number (maxDPBsize) of the reconstructed pictures storable in the frame memory equal to a doubled MFSBuffer, and

wherein a number of the reconstructed pictures stored in the frame memory is no more than a number set in the maxDPBsize.

2. The video predictive encoding device according to claim 1, wherein the memory

manager is further executable with the programmed CPU to:

compare a quartered MaxLumaFS with the target picture frame size;

upon a determination that the quartered MaxLumaFS is equal to or larger than the target picture frame size, quadruple the MFSBuffer; and

set the maxDPBsize equal to a quadrupled MFSBuffer.

3. A video predictive decoding device comprising:

a CPU programmed to implement video predictive decoding;

a receiver executable with the CPU to receive encoded data comprising compressed picture data representative of a compressed form of a plurality of pictures forming a video sequence, wherein the plurality of pictures are encoded by either intra-frame prediction or inter-frame prediction;

a decoder executable with the CPU to decode the compressed picture data to reconstruct the pictures;

a decoded picture buffer (DPB) executable with the CPU to store one or more of the reconstructed pictures as reference pictures to be used for decoding a subsequent picture; and

a buffer manager executable with the programmed CPU to manage the DPB, the buffer manager being executable to retrieve (i) a target picture frame size indicative of a frame size of a target picture, (ii) a maximum frame size (MaxLumaFS) defined in advance by level information and indicative of a maximum size of a reconstructed picture storable in the DPB, and (iii) a size of the frame memory (MFSBuffer) expressed by a preset maximum number of reconstructed pictures storable in the DPB,

wherein the buffer manager is executable to:

compare a halved MaxLumaFS with the target picture frame size;

upon a determination that the halved MaxLumaFS is equal to or larger than the target picture frame size, double the MFSBuffer; and

set an adaptive maximum number (maxDPBsize) of reconstructed pictures storable in the DPB equal to a doubled MFSBuffer, and

wherein a number of the reconstructed pictures stored in the DPB is no more than a number set in the maxDPBsize.

4. The video predictive decoding device according to claim 3, wherein the buffer manager is further executable to:

compare a quartered MaxLumaFS with the target picture frame size;

upon a determination that the quartered MaxLumaFS is equal to or larger than the target picture frame size, quadruple the MFSBuffer; and

set the maxDPBsize equal to a quadrupled MFSBuffer.

5. A video predictive encoding method executed by a video predictive encoding device, comprising:

receiving a plurality of pictures forming a video sequence;

encoding the received pictures by either intra-frame prediction or inter-frame prediction to generate compressed picture data;

decoding the compressed picture data to reconstruct the pictures;

storing one or more of the reconstructed pictures in a frame memory as reference pictures to be used for encoding a subsequent picture;

retrieving (i) a target picture frame size indicative of a frame size of a target picture, (ii) a maximum frame size (MaxLumaFS) defined in advance by level information and indicative of a maximum size of a reconstructed picture storable in the frame memory, and (iii) a size of the frame memory (MFSBuffer) expressed by a preset maximum number of reconstructed pictures storable in the frame memory;

comparing a halved MaxLumaFS with the target picture frame size;

upon a determination that the halved MaxLumaFS is equal to or larger than the target picture frame size, doubling the MFSBuffer; and

setting an adaptive maximum number (maxDPBsize) of the reconstructed pictures storable in the frame memory equal to a doubled MFSBuffer,

wherein a number of the reconstructed pictures stored in the frame memory is no more than a number set in the maxDPBsize.

6. The video predictive encoding method according to claim 5, further comprising:

comparing a quartered MaxLumaFS with the target picture frame size;

upon a determination that the quartered MaxLumaFS is equal to or larger than the target picture frame size, quadrupling the MFSBuffer; and

setting the maxDPBsize equal to a quadrupled MFSBuffer.

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

receiving encoded data comprising compressed picture data representative of a compressed form of a plurality of pictures forming a video sequence, wherein the plurality of pictures are encoded by either intra-frame prediction or inter-frame prediction;

decoding the compressed form of the plurality of pictures to reconstruct the plurality of pictures;

storing in a decoded picture buffer (DPB) one or more of the reconstructed pictures as reference pictures to be used for decoding a subsequent picture;

retrieving (i) a target picture frame size indicative of a frame size of a target picture, (ii) a maximum frame size (MaxLumaFS) defined in advance by level information and indicative of a maximum size of a reconstructed picture storable in the DPB, and (iii) a size of the frame memory (MFSBuffer) expressed by a preset maximum number of reconstructed pictures storable in the DPB;

comparing a halved MaxLumaFS with the target picture frame size;

upon a determination that the halved MaxLumaFS is equal to or larger than the target picture frame size, doubling the MFSBuffer; and

setting an adaptive maximum number (maxDPBsize) of the reconstructed pictures storable in the DPB equal to a doubled MFSBuffer,

wherein a number of the reconstructed pictures stored in the DPB is no more than a number set in the maxDPBsize.

8. The video predictive decoding method according to claim 7, further comprising:

comparing a quartered MaxLumaFS with the target picture frame size;

upon a determination that the quartered MaxLumaFS is equal to or larger than the target picture frame size, quadrupling the MFSBuffer; and

setting the maxDPBsize equal to a quadrupled MFSBuffer.

9. The video predictive decoding method according to claim 7, further comprising leaving memory pointers kept pointing storage locations of the DPB partitioned according to the MFSBuffer, after setting the maxDPBsize equal to the doubled MFSBuffer.

10. The video predictive decoding method according to claim 8, further comprising leaving memory pointers kept pointing storage locations of the DPB partitioned according to the MFSBuffer, after setting the maxDPBsize equal to the quadrupled MFSBuffer.

11. A non-transitory storage medium that stores a video predictive decoding program which programs a CPU to execute:

a receiver operable to receive encoded data comprising compressed picture data representative of a compressed form of a plurality of pictures forming a video sequence, wherein the plurality of pictures are encoded by either intra-frame prediction or inter-frame prediction;

a decoder operable to decode the compressed picture data to reconstruct the pictures;

a decoded picture buffer (DPB) operable to store one or more of the reconstructed pictures as reference pictures to be used for decoding a subsequent picture; and

a buffer manager operable to manage the DPB, the buffer manager being executable to retrieve (i) a target picture frame size indicative of a frame size of a target picture, (ii) a maximum frame size (MaxLumaFS) defined in advance by level information and indicative of a

maximum size of a reconstructed picture storable in the DPB, and (iii) a size of the frame memory (MFSBuffer) expressed by a preset maximum number of reconstructed pictures storable in the DPB; and

wherein the buffer manager further operable to:

compare a halved MaxLumaFS with the target picture frame size;

upon a determination that the halved MaxLumaFS is equal to or larger than the target picture frame size, double the MFSBuffer; and

set an adaptive maximum number (maxDPBsize) of the reconstructed pictures storable in the DPB equal to a doubled MFSBuffer, and

wherein a number of the reconstructed pictures stored in the DPB is no more than a number set in the maxDPBsize.

NTT DOCOMO, INC.

Patent Attorneys for the Applicant

SPRUSON & FERGUSON

Fig.1

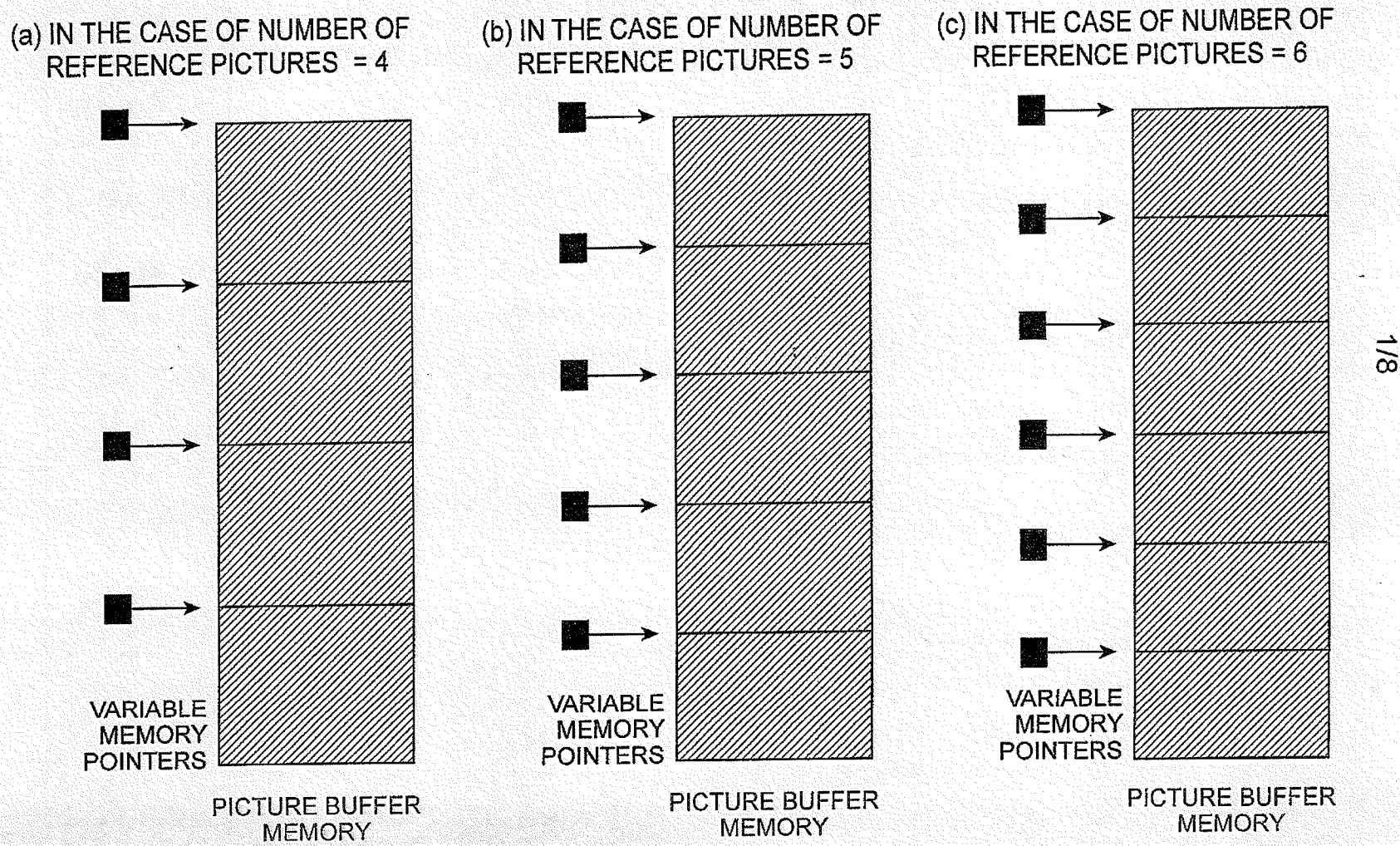

(a) IN THE CASE OF NUMBER OF REFERENCE PICTURES = 4
VARIABLE MEMORY POINTERS
PICTURE BUFFER MEMORY
(b) IN THE CASE OF NUMBER OF REFERENCE PICTURES = 5
VARIABLE MEMORY POINTERS
PICTURE BUFFER MEMORY
(c) IN THE CASE OF NUMBER OF REFERENCE PICTURES = 6
VARIABLE MEMORY POINTERS
PICTURE BUFFER MEMORY

Fig.2

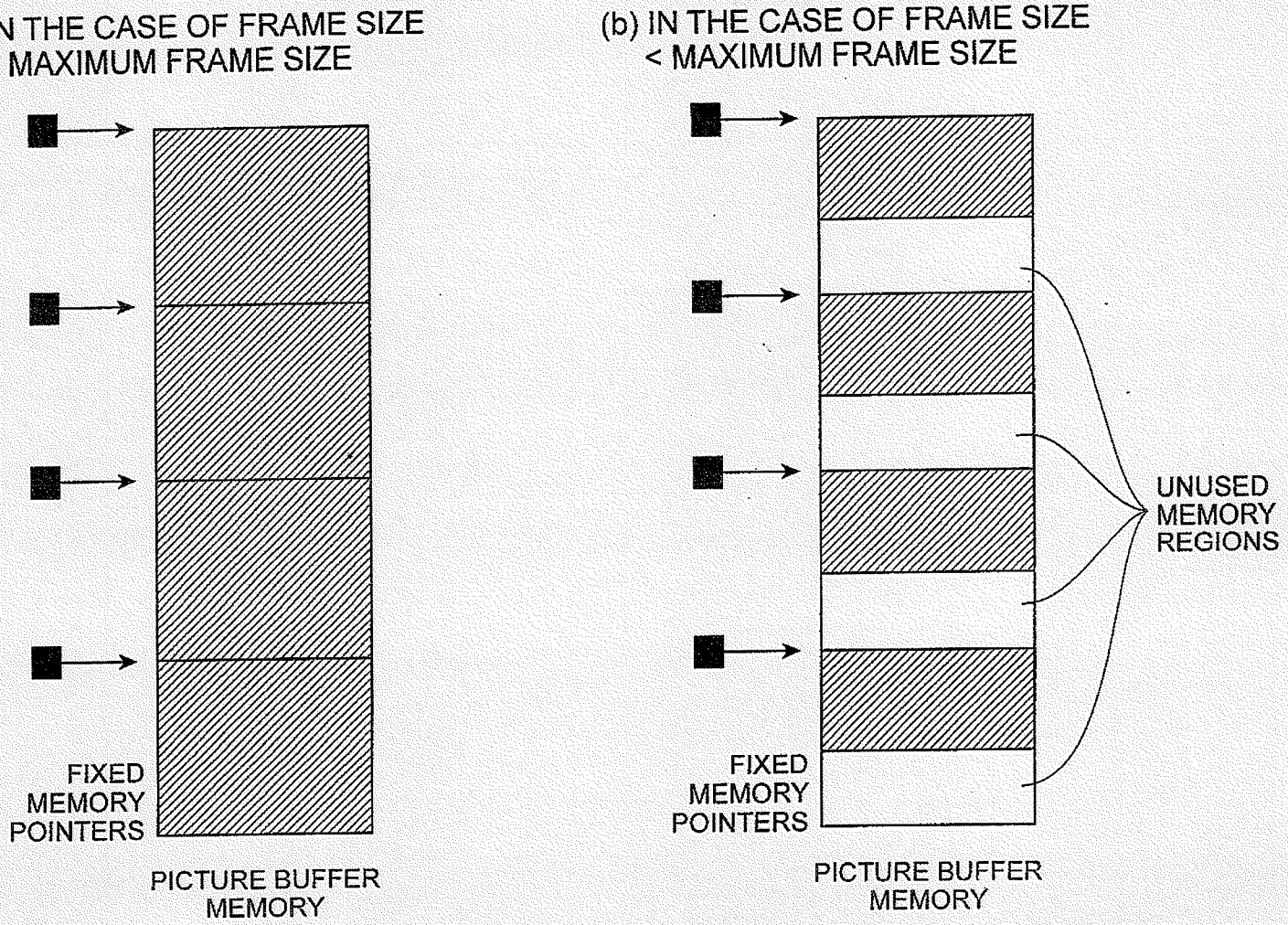

(a) IN THE CASE OF FRAME SIZE
= MAXIMUM FRAME SIZE
FIXED
MEMORY
POINTERS
PICTURE BUFFER
MEMORY
(b) IN THE CASE OF FRAME SIZE
< MAXIMUM FRAME SIZE
UNUSED
MEMORY
REGIONS
FIXED
MEMORY
POINTERS
PICTURE BUFFER
MEMORY

Fig.3

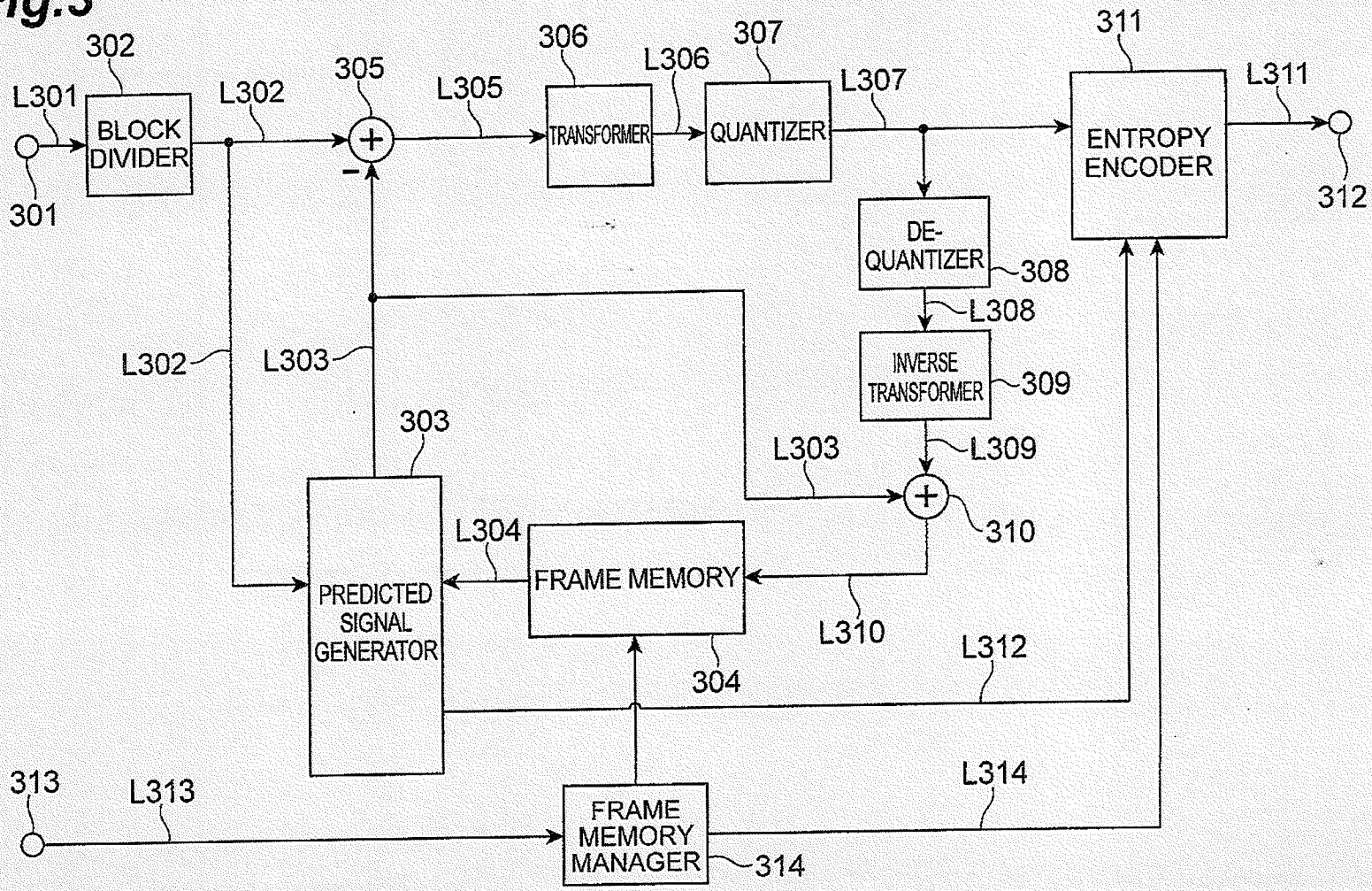

302
L301
BLOCK DIVIDER
L302
305
L305
306
TRANSFORMER
L306
307
QUANTIZER
L307
311
ENTROPY ENCODER
L311
301
312
DE-QUANTIZER
308
L308
INVERSE TRANSFORMER
309
L309
L303
303
310
L304
FRAME MEMORY
PREDICTED SIGNAL GENERATOR
L310
L312
304
313
L313
FRAME MEMORY MANAGER
314
L314

Fig.4

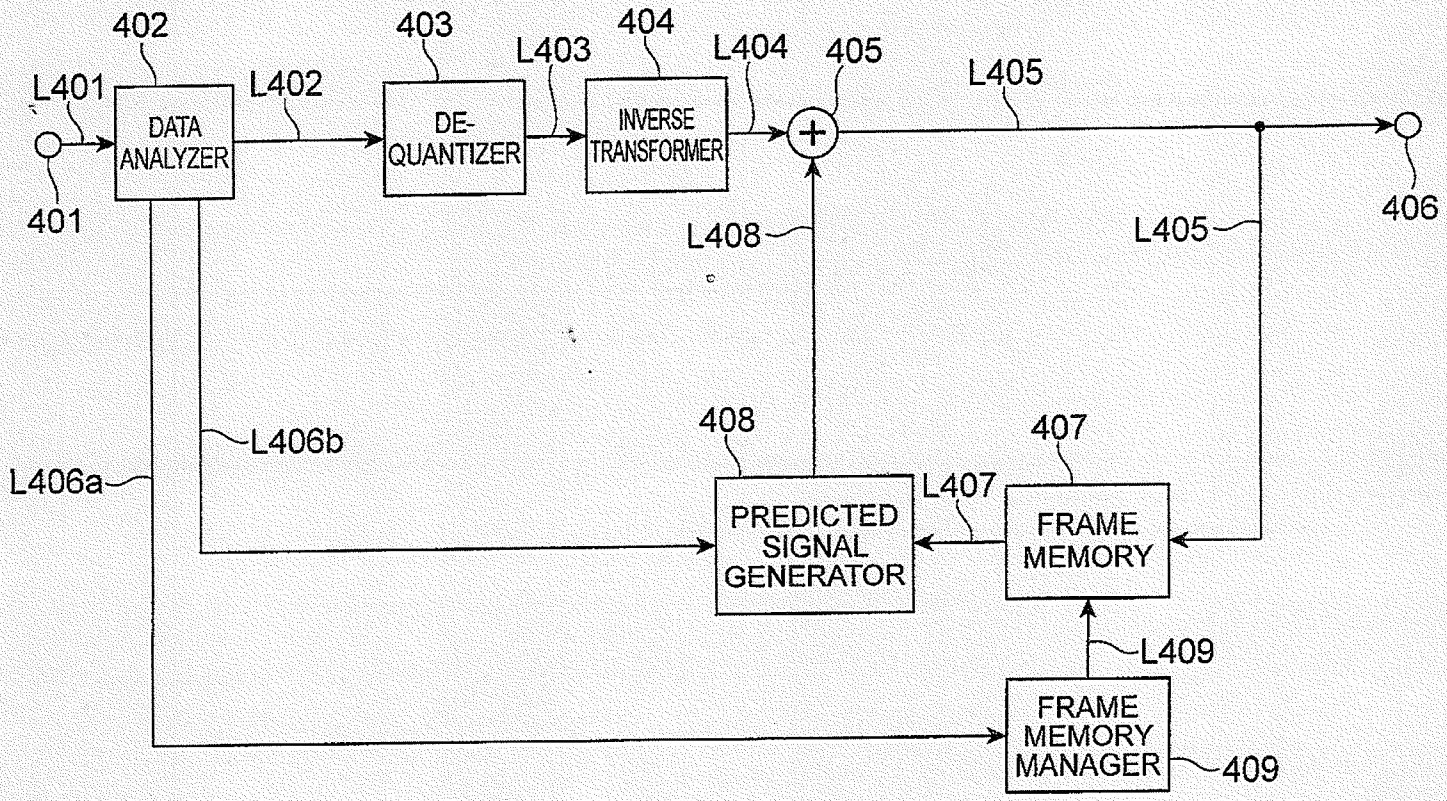

402
403
404
L403
L404
405
L401
L402
L405
DATA ANALYZER
DE-QUANTIZER
INVERSE TRANSFORMER
401
406
L408
L405
L406b
408
407
L406a
L407
PREDICTED SIGNAL GENERATOR
FRAME MEMORY
L409
FRAME MEMORY MANAGER
409

Fig.5

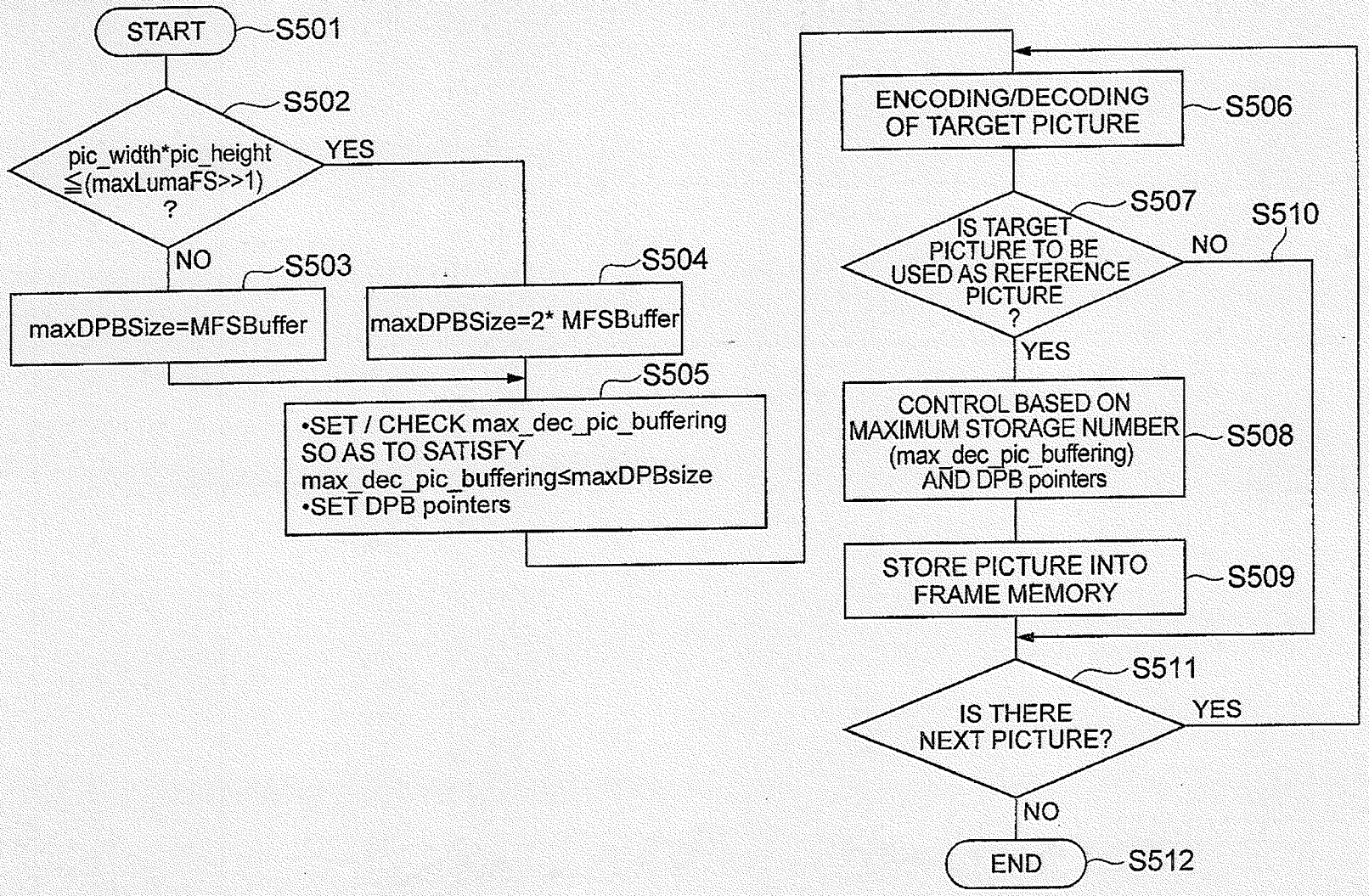

START
S501
S502
pic_width*pic_height ≦(maxLumaFS>>1) ?
YES
NO
S503
maxDPBSize=MFSBuffer
S504
maxDPBSize=2* MFSBuffer
S505
•SET / CHECK max_dec_pic_buffering SO AS TO SATISFY max_dec_pic_buffering≤maxDPBsize
•SET DPB pointers
ENCODING/DECODING OF TARGET PICTURE
S506
S507
IS TARGET PICTURE TO BE USED AS REFERENCE PICTURE ?
S510
NO
YES
CONTROL BASED ON MAXIMUM STORAGE NUMBER (max_dec_pic_buffering) AND DPB pointers
S508
STORE PICTURE INTO FRAME MEMORY
S509
S511
IS THERE NEXT PICTURE?
YES
NO
END
S512

Fig.6

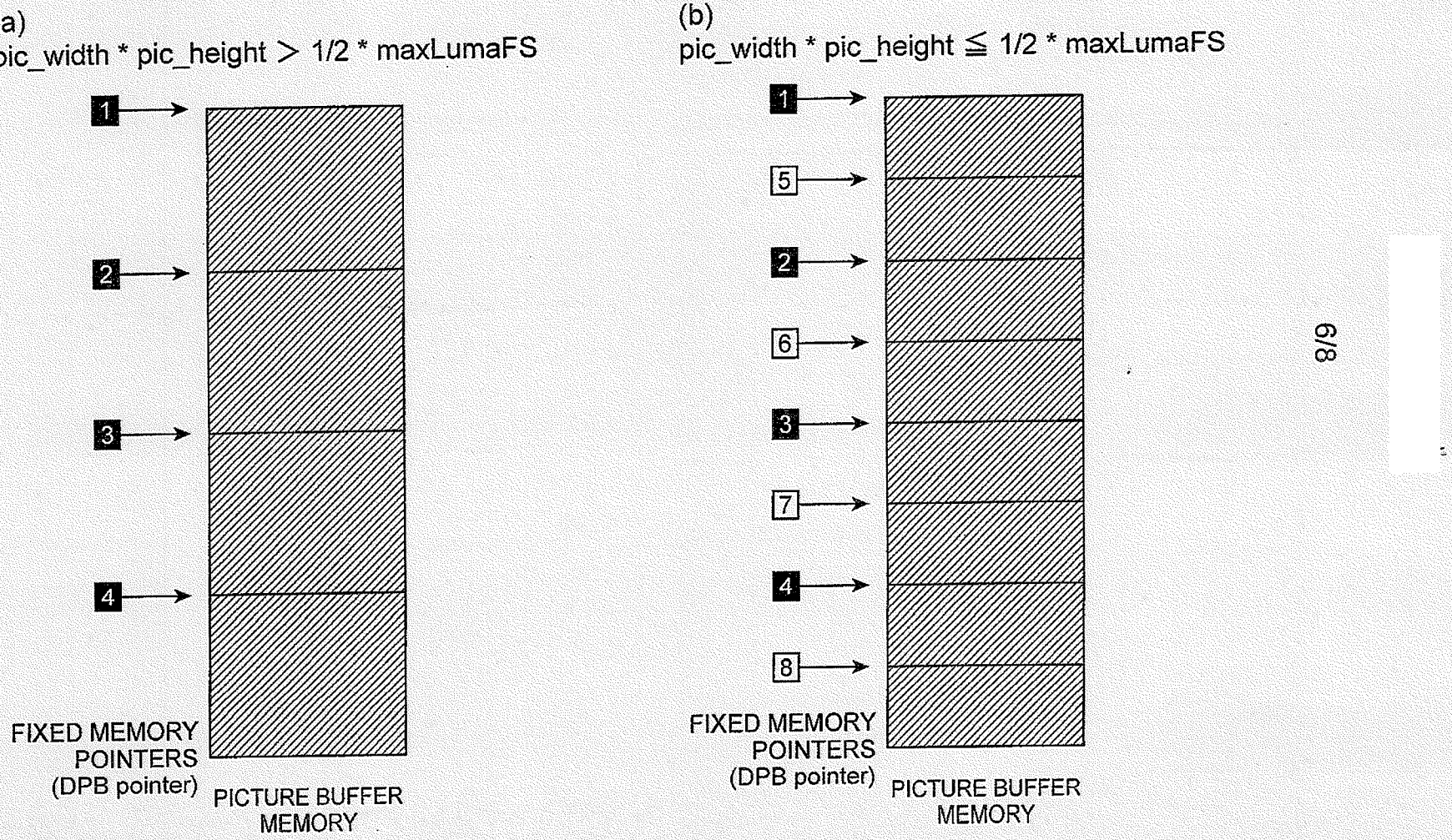
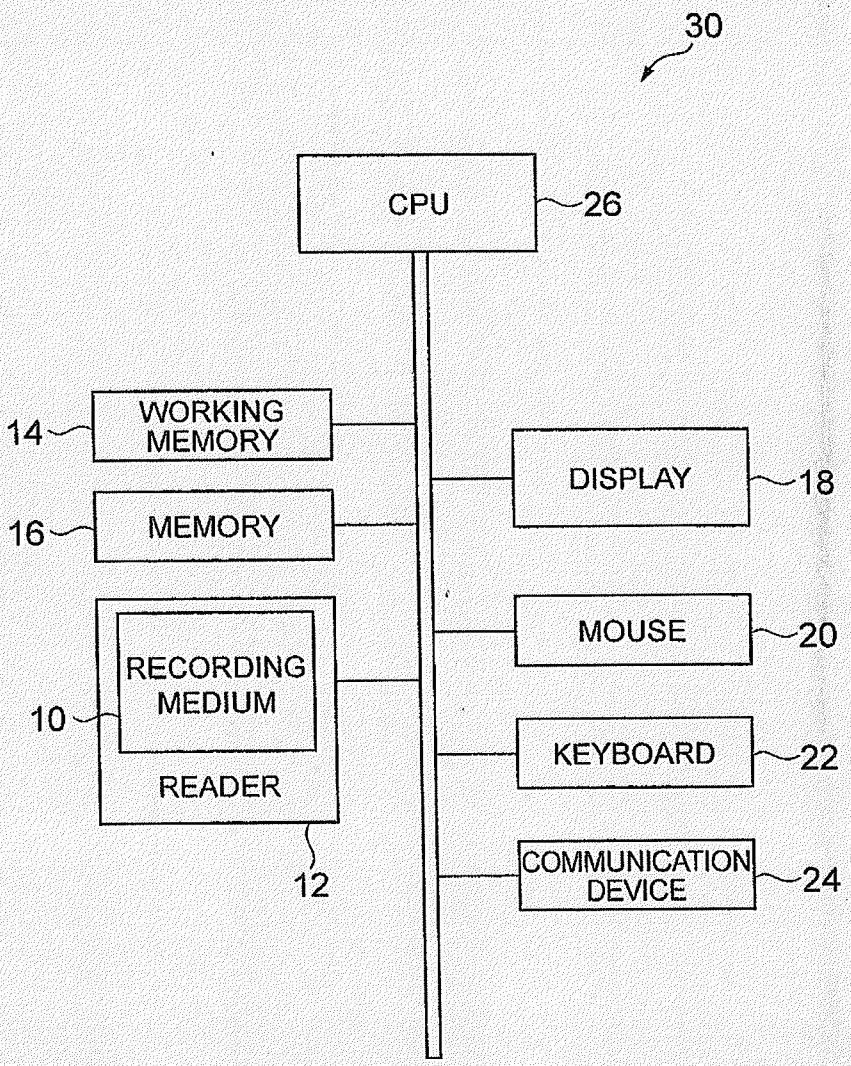

(a)
pic_width * pic_height > 1/2 * maxLumaFS
1
2
3
4
FIXED MEMORY POINTERS (DPB pointer)
PICTURE BUFFER MEMORY
(b)
pic_width * pic_height ≦ 1/2 * maxLumaFS
1
5
2
6
3
7
4
8
FIXED MEMORY POINTERS (DPB pointer)
PICTURE BUFFER MEMORY

Fig.7



2020201376 25 Feb 2020

Fig.8