012506 17224 U.S. PTO

113260 U.S.PTO 11/338934 012506

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant: HANNUKSELA

Docket: TPO.004.C2

Title: GROUPING OF IMAGE FRAMES IN VIDEO CODING

CERTIFICATE UNDER 37 CFR 1.10:
"Express Mail" mailing label number: EV 712609393 US
Date of Deposit: January 25, 2006
I hereby certify that this paper or fee is being deposited with the U.S. Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above and is addressed to Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

By: Tracey M. Dotter
Tracey M. Dotter

CONTINUATION PATENT APPLICATION UNDER 37 C.F.R. § 1.53(b)

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

This is a request for filing a continuation patent application under 37 CFR § 1.53(b) of pending patent application serial number 10/306,942 filed on November 29, 2002, which claims priority under 35 U.S.C. §119 from corresponding Finnish patent application No. 20020127, filed in Finland on January 23, 2002.

We are transmitting herewith the following:

- ☒ Transmittal sheet containing Certificate under 37 CFR 1.10 (2 pages).
- ☒ Patent Application: Pages numbered 1-38, with 1pg. Abstract; 16 claims.
- ☒ 4 sheets of formal drawings, having Figures 1-2, 3A-3B, 4A-4C, 5A-5C, 6A-6C, and 7.
- ☒ Copy of the executed Declaration from the priority application (3 pages).
- ☒ A copy of the Change of Correspondence Address submitted in the prior pending application. Please send all future correspondence to the following address:

HOLLINGSWORTH & FUNK, LLC
8009 34th Avenue South, Suite 125
Minneapolis, MN 55425
952.854.2700

- ☒ The priority pending application is assigned to: Nokia Corporation at Reel/Frame No.: 013741/0536.

AMAZON - 1002
IPR2024-00691

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant: HANNUKSELA

Docket: TPO.004.C1

Title: GROUPING OF IMAGE FRAMES IN VIDEO CODING

☒ The Power of Attorney in the prior pending application is to: Steven R. Funk, Reg. No. 37,830.

☒ Authorization is given to charge Deposit Account No. 50-3581 (TPO.004.C2) in the amount of $1,800.00 for the filing fees. Authorization is also given to charge/credit Deposit Account No. 50-3581 (TPO.004.C2) for any deficiencies/overpayments.

☒ 1 Return Postcard

The fees are calculated as follows:

Basic Filing Fee =	$ 300.00
Search Fee =	$ 500.00
Examination Fee =	$ 200.00
Fee for total # of claims over 20 = $50 * (16 – 20) =	$ 0.00
Fee for total # of independent claims over 3 = $200 * (7 – 3) =	$ 800.00
TOTAL FILING FEES DUE =	$1,800.00

Respectfully submitted,

HOLLINGSWORTH & FUNK, LLC
8009 34th Avenue South, Suite 125
Minneapolis, MN 55425
952.854.2700

Date: January 25, 2006

By: _____
Steven R. Funk
Reg. No. 37,830

OIPE JC48
AUG 17 2005
PATENT & TRADEMARK OFFICE

PATENT APPLICATION

IN THE UNITED STATES PATENT & TRADEMARK OFFICE

Applicant:	Miska Hannuksela	Examiner:	Unassigned
Serial No.:	10/306,942	Group Art Unit:	2125
Filed:	11/29/2002	**New Docket No.:**	**TAPA.004PA**
		Old Docket No.:	KOLS.021PA

CERTIFICATE UNDER 37 CFR 1.8: The undersigne d hereby certifies that this correspondence and the papers, as described hereinabove, are being deposited in the United States Postal Service, as first class mail, in an envelope addressed to: Commissioner for Patents, United States Patent and Trademark O ffice, P.O. Box 1450, Alexandria, VA 22313-1450, on August 15, 2005.

By: _____

CHANGE OF CORRESPONDENCE ADDRESS

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir or Madam:

Please update the attorney docket number as reflected above (**TAPA.004PA**), and change the correspondence Address for the above-referenced application as follows:

Hollingsworth & Funk, LLC
8009 34th Avenue South, Suite 125
Minneapolis, MN 55425
(952) 854-2700

If there are any questions regarding this matter, please call the undersigned at 952-854-2700.

Respectfully submitted,

By: _____

Steven R. Funk
Reg. No.: 37,830

17224 U.S. PTO
012506

113260 U.S.PTO
11/338934
012506

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant: HANNUKSELA

Docket: TPO.004.C2

Title: GROUPING OF IMAGE FRAMES IN VIDEO CODING

CERTIFICATE UNDER 37 CFR 1.10:
"Express Mail" mailing label number: EV 712609393 US
Date of Deposit: January 25, 2006
I hereby certify that this paper or fee is being deposited with the U.S. Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above and is addressed to Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

By: Tracey M. Dotter
Tracey M. Dotter

CONTINUATION PATENT APPLICATION UNDER 37 C.F.R. § 1.53(b)

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

This is a request for filing a continuation patent application under 37 CFR § 1.53(b) of pending patent application serial number 10/306,942 filed on November 29, 2002, which claims priority under 35 U.S.C. §119 from corresponding Finnish patent application No. 20020127, filed in Finland on January 23, 2002.

We are transmitting herewith the following:

☒ Transmittal sheet containing Certificate under 37 CFR 1.10 (2 pages).
☒ Patent Application: Pages numbered 1-38, with 1pg. Abstract; 16 claims.
☒ 4 sheets of formal drawings, having Figures 1-2, 3A-3B, 4A-4C, 5A-5C, 6A-6C, and 7.
☒ Copy of the executed Declaration from the priority application (3 pages).
☒ A copy of the Change of Correspondence Address submitted in the prior pending application. Please send all future correspondence to the following address:

HOLLINGSWORTH & FUNK, LLC
8009 34th Avenue South, Suite 125
Minneapolis, MN 55425
952.854.2700

☒ The priority pending application is assigned to: Nokia Corporation at Reel/Frame No.: 013741/0536.

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant: HANNUKSELA

Docket: TPO.004.C1

Title: GROUPING OF IMAGE FRAMES IN VIDEO CODING

☒ The Power of Attorney in the prior pending application is to: Steven R. Funk, Reg. No. 37,830.

☒ Authorization is given to charge Deposit Account No. 50-3581 (TPO.004.C2) in the amount of $1,800.00 for the filing fees. Authorization is also given to charge/credit Deposit Account No. 50-3581 (TPO.004.C2) for any deficiencies/overpayments.

☒ 1 Return Postcard

The fees are calculated as follows:

Basic Filing Fee =	$ 300.00
Search Fee =	$ 500.00
Examination Fee =	$ 200.00
Fee for total # of claims over 20 = $50 * (16 – 20) =	$ 0.00
Fee for total # of independent claims over 3 = $200 * (7 – 3) =	$ 800.00
TOTAL FILING FEES DUE =	$1,800.00

Respectfully submitted,

HOLLINGSWORTH & FUNK, LLC
8009 34th Avenue South, Suite 125
Minneapolis, MN 55425
952.854.2700

Date: January 25, 2006

By: _____
Steven R. Funk
Reg. No. 37,830

OIPE JC48
AUG 17 2005
PATENT & TRADEMARK OFFICE

PATENT APPLICATION

IN THE UNITED STATES PATENT & TRADEMARK OFFICE

Applicant:	Miska Hannuksela	Examiner:	Unassigned
Serial No.:	10/306,942	Group Art Unit:	2125
Filed:	11/29/2002	**New Docket No.:**	**TAPA.004PA**
		Old Docket No.:	KOLS.021PA

CERTIFICATE UNDER 37 CFR 1.8: The undersigne d hereby certifies that this correspondence and the papers, as described hereinabove, are being deposited in the United States Postal Service, as first class mail, in an envelope addressed to: Commissioner for Patents, United States Patent and Trademark O ffice, P.O. Box 1450, Alexandria, VA 22313-1450, on August 15, 2005.

By: _____

CHANGE OF CORRESPONDENCE ADDRESS

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir or Madam:

Please update the attorney docket number as reflected above (**TAPA.004PA**), and change the correspondence Address for the above-referenced application as follows:

Hollingsworth & Funk, LLC
8009 34th Avenue South, Suite 125
Minneapolis, MN 55425
(952) 854-2700

If there are any questions regarding this matter, please call the undersigned at 952-854-2700.

Respectfully submitted,

By: _____

Steven R. Funk
Reg. No.: 37,830

GROUPING OF IMAGE FRAMES IN VIDEO CODING

RELATED PATENT DOCUMENTS

[0001] This application is a continuation of U.S. Patent Application Serial No. 10/306,942 filed on November 29, 2002 which is incorporated herein by reference in its entirety.

FIELD OF THE INVENTION

[0002] The invention relates to the grouping of multimedia files, particularly video files and particularly in connection with streaming.

BACKGROUND OF THE INVENTION

[0003] The term 'streaming' refers to simultaneous sending and playback of data, typically multimedia data, such as audio and video files, in which the recipient may begin data playback already before all the data to be transmitted has been received. Multimedia data streaming systems comprise a streaming server and terminal devices that the recipients use for setting up a data connection, typically via a telecommunications network, to the streaming server. From the streaming server the recipients retrieve either stored or real-time multimedia data, and the playback of the multimedia data can then begin, most advantageously almost in real-time with the transmission of the data, by means of a streaming application included in the terminal.

[0004] From the point of view of the streaming server, the streaming may be carried out either as normal streaming or as progressive downloading to the terminal. In normal streaming the transmission of the multimedia data and/or the data contents are controlled either by making sure that the bit rate of the transmission substantially corresponds to the playback rate of the terminal device, or, if the telecommunications network used in the transmission causes a bottleneck in data transfer, by making sure that the bit rate of the transmission substantially corresponds to the bandwidth available in the telecommunications network. In progressive downloading the transmission of the multimedia data and/or the data contents do not necessarily have to be interfered with at all, but the multimedia files are transmitted as such to the recipient, typically by using transfer protocol flow control. The terminals then receive, store and reproduce an exact copy of the data transmitted from the server, which copy can then be later reproduced again on the terminal without

needing to start a streaming again via the telecommunications network. The multimedia files stored in the terminal are, however, typically very large and their transfer to the terminal is time-consuming, and they require a significant amount of storage memory capacity, which is why a normal streaming is often preferred.

[0005] The video files in multimedia files comprise a great number of still image frames, which are displayed rapidly in succession (of typically 15 to 30 frames per s) to create an impression of a moving image. The image frames typically comprise a number of stationary background objects, determined by image information which remains substantially unchanged, and few moving objects, determined by image information that changes to some extent. The information comprised by consecutively displayed image frames is typically largely similar, i.e. successive image frames comprise a considerable amount of redundancy. The redundancy appearing in video files can be divided into spatial, temporal and spectral redundancy. Spatial redundancy refers to the mutual correlation of adjacent image pixels, temporal redundancy refers to the changes taking place in specific image objects in subsequent frames, and spectral redundancy to the correlation of different colour components within an image frame.

[0006] To reduce the amount of data in video files, the image data can be compressed into a smaller form by reducing the amount of redundant information in the image frames. In addition, while encoding, most of the currently used video encoders downgrade image quality in image frame sections that are less important in the video information. Further, many video coding methods allow redundancy in a bit stream coded from image data to be reduced by efficient, lossless coding of compression parameters known as VLC (Variable Length Coding).

[0007] In addition, many video coding methods make use of the above-described temporal redundancy of successive image frames. In that case a method known as motion-compensated temporal prediction is used, i.e. the contents of some (typically most) of the image frames in a video sequence are predicted from other frames in the sequence by tracking changes in specific objects or areas in successive image frames. A video sequence always comprises some compressed image frames the image information of which has not been determined using motion-compensated temporal prediction. Such frames are called INTRA-frames, or I-frames.

Correspondingly, motion-compensated video sequence image frames predicted from previous image frames, are called INTER-frames, or P-frames (Predicted). The image information of P-frames is determined using one I-frame and possibly one or more previously coded P-frames. If a frame is lost, frames dependent on it can no longer be correctly decoded.

[0008] An I-frame typically initiates a video sequence defined as a Group of Pictures (GOP), the P-frames of which can only be determined on the basis of the I-frame and the previous P-frames of the GOP in question. The next I-frame begins a new group of pictures GOP, the image information comprised by which cannot thus be determined on the basis of the frames of the previous GOP. In other words, groups of pictures are not temporally overlapping, and each group of picture can be decoded separately. In addition, many video compression methods employ bi-directionally predicted B-frames (Bi-directional), which are set between two anchor frames (I- and P-frames, or two P-frames) within a group of pictures GOP, the image information of a B-frame being predicted from both the previous anchor frame and the one succeeding the B-frame. B-frames therefore provide image information of higher quality than P-frames, but typically they are not used as anchor frames, and therefore their removal from the video sequence does not degrade the quality of subsequent images. However, nothing prevents B-frames from being used as anchor frames as well, only in that case they cannot be removed from the video sequence without deteriorating the quality of the frames dependent on them.

[0009] Each video frame may be divided into what are known as macroblocks that comprise the colour components (such as Y, U, V) of all pixels of a rectangular image area. More specifically, a macroblock consists of at least one block per colour component, the blocks each comprising colour values (such as Y, U or V) of one colour level in the image area concerned. The spatial resolution of the blocks may differ from that of the macroblocks, for example U- and V-components may be displayed using only half of the resolution of Y-component. Macroblocks can be further grouped into slices, for example, which are groups of macroblocks that are typically selected in the scanning order of the image. Temporal prediction is typically carried out in video coding methods block- or macroblock-specifically, instead of image-frame-specifically.

[0010] To allow for flexible streaming of video files, many video coding systems employ scalable coding in which some elements or element groups of a video sequence can be removed without affecting the reconstruction of other parts of the video sequence. Scalability is typically implemented by grouping the image frames into a number of hierarchical layers. The image frames coded into the image frames of the base layer substantially comprise only the ones that are compulsory for the decoding of the video information at the receiving end. The base layer of each group of pictures GOP thus comprises one I-frame and a necessary number of P-frames. One or more enhancement layers can be determined below the base layer, each one of the layers improving the quality of the video coding in comparison with an upper layer. The enhancement layers thus comprise P- or B-frames predicted on the basis of motion-compensation from one or more upper layer images. The frames are typically numbered according to an arithmetical series.

[0011] In streaming, transmission bit rate must be controllable either on the basis of the bandwidth to be used or the maximum decoding or bit rate value of the recipient. Bit rate can be controlled either at the streaming server or in some element of the telecommunications network, such as an Internet router or a base station of a mobile communications network. The simplest means for the streaming server to control the bit rate is to leave out B-frames having a high information content from the transmission. Further, the streaming server may determine the number of scalability layers to be transmitted in a video stream, and thus the number of the scalability layers can be changed always when a new group of pictures GOP begins. It is also possible to use different video sequence coding methods. Correspondingly, B-frames, as well as other P-frames of the enhancement layers, can be removed from the bit stream in a telecommunications network element.

[0012] The above arrangement involves a number of drawbacks. Many coding methods, such as the coding according to the ITU-T (International Telecommunications Union, Telecommunications Standardization Sector) standard H.263, are familiar with a procedure called reference picture selection. In reference picture selection at least a part of a P-image has been predicted from at least one other image than the one immediately preceding the P-image in the time domain. The selected reference image is signalled in a coded bit stream or in bit stream header fields image-,

image-segment- (such as a slice or a group of macroblocks), macroblock-, or block-specifically. The reference picture selection can be generalized such that the prediction can also be made from images temporally succeeding the image to be coded. Further, the reference picture selection can be generalized to cover all temporally predicted frame types, including B-frames. Since it is possible to also select at least one image preceding an I-image that begins a group of pictures GOP as the reference image, a group of pictures employing reference picture selection cannot necessarily be decoded independently. In addition, the adjusting of scalability or coding method in the streaming server or a network element becomes difficult, because the video sequence must be decoded, parsed and buffered for a long period of time to allow any dependencies between different image groups to be detected.

[0013] A further problem relates to detection of image frames from which a decoder can start the decoding process. The detection is useful for multiple purposes. For example, an end-user may wish to start browsing a video file from the middle of a video sequence. Another example relates to starting the reception of a broadcast or multicast video transmission from the middle of the video transmission. A third example relates to on-demand streaming from a server and occurs when an end-user wishes to start playback from a certain position of a stream.

BRIEF DESCRIPTION OF THE INVENTION

[0014] Now there is invented an improved method and equipment implementing the method, that enable detection of image frames from which a decoder can start the decoding process. Various aspects of the invention include methods, a video encoder, a video decoder, and computer programs that are characterized by what is stated in the independent claims. The preferred embodiments of the invention are disclosed in the dependent claims.

[0015] The invention is based on the idea of encoding a video sequence comprising an independent sequence of image frames, wherein at least one reference picture is predictable from at least one previous image frame that is earlier than the previous reference image frame in decoding order. An indication of at least one image frame is encoded into the video sequence, which indicated image frame is the first picture, in decoding order, of the independent sequence, said at least one reference image frame being included in the sequence. Respectively, in the decoding phase, the indication

of at least one image frame is decoded from the video sequence, and the decoding of the video sequence is started from said first image frame of the independent sequence, whereby the video sequence is decoded without prediction from any image frame decoded prior to said first image frame.

[0016] Consequently, the idea of the present invention is to determine an initiation picture in an independently decodable group of pictures, whereby in the decoding phase any picture prior to said initiation picture is defined as unusable for reference. Accordingly, after decoding the initiation picture, all following pictures of the independently decodable sequence can be decoded without prediction from any picture decoded prior to said initiation picture.

[0017] According to an embodiment, the indication is encoded into the video sequence as a separate flag included in the header of a slice.

[0018] According to an embodiment, identifier values for pictures are encoded according to a numbering scheme, and the identifier value for the indicated first picture of an independent sequence is reset, preferably to zero.

[0019] According to an embodiment, an identifier value for the independent sequence is encoded into the video sequence.

[0020] An advantage of the procedure of the invention is that it enables to start browsing a video sequence from a random point, i.e. it provides the decoder with the information about the first picture of an independently decodable sequence. Thus, the decoder knows that by decoding this first picture, it is possible to continue the decoding process without any prediction from any prior picture. Accordingly, a further advantage is that the decoder may discard any picture decoded prior to said initiation picture from its buffer memory, since the prior pictures are no longer needed in the decoding process. A further advantage is that the procedure of the invention allows a separate video sequence to the easily inserted into another video sequence.

[0021] A further advantage is that enables identification of a picture boundary between two back-to-back initiation pictures by looking at the sub-sequence number of the initiation pictures. A yet further advantage relates to detection of losses of image frames that start an independently decodable sub-sequence. If such an image frame is lost e.g. during transmission, it is unlikely that any error concealment method would result into a subjectively satisfactory image quality. It is therefore an advantage that decoders are provided means

to detect losses of image frames that start an independently decodable sub-sequence. Decoders may react to such a loss by requesting a retransmission or picture refresh, for example.

BRIEF DESCRIPTION OF THE DRAWINGS

[0022] In the following, the invention will be described in connection with the preferred embodiments and with reference to the accompanying drawings, in which

Figure 1 illustrates a common multimedia data streaming system in which the scalable coding hierarchy of the invention can be applied;

Figure 2 illustrates a scalable coding hierarchy of a preferred embodiment of the invention;

Figures 3a and 3b illustrate embodiments of the invention for adjusting scalability;

Figures 4a, 4b and 4c illustrate embodiments of the invention for adjusting image numbering;

Figures 5a, 5b and 5c illustrate embodiments of the invention for using B-frames in a scalable coding hierarchy;

Figures 6a, 6b and 6c illustrate scalable coding hierarchies of preferred embodiments of the invention in connection with reference picture selection; and

Figure 7 illustrates an arrangement according to a preferred embodiment of the invention for coding scene transition.

DETAILED DESCRIPTION OF THE INVENTION

[0023] In the following, a general-purpose multimedia data streaming system is disclosed, the basic principles of which can be applied in connection with any telecommunications system. Although the invention is described here with a particular reference to a streaming system, in which the multimedia data is transmitted most preferably through a telecommunications network employing a packet-switched data protocol, such as an IP network, the invention can equally well be implemented in circuit-switched networks, such as fixed telephone networks PSTN/ISDN (Public Switched Telephone Network/Integrated Services Digital Network) or in mobile communications networks PLMN (Public Land Mobile Network). Further, the invention can be applied in the streaming of multimedia files in the form of both normal

streaming and progressive downloading, and for implementing video calls, for example.

[0024] It is also to be noted that although the invention is described here with a particular reference to streaming systems and the invention can also be advantageously applied in them, the invention is not restricted to streaming systems alone, but can be applied in any video reproduction system, irrespective of how a video file that is to be decoded is downloaded and where it is downloaded from. The invention can therefore be applied for example in the playback of a video file to be downloaded from a DVD disc or from some other computer memory carrier, for example in connection with varying processing capacity available for video playback. In particular, the invention can be applied to different video codings of low bit rate that are typically used in telecommunications systems subject to bandwidth restrictions. One example is the system defined in the ITU-T standard H.263 and the one that is being defined in H.26L (possibly later to become H.264). In connection with these, the invention can be applied to mobile stations, for example, in which case the video playback can be made to adjust both to changing transfer capacity or channel quality and to the processor power currently available, when the mobile station is used also for executing other applications than video playback.

[0025] It is further to be noted that, for the sake of clarity, the invention will be described below by giving an account of image frame coding and temporal predicting on image frame level. However, in practice coding and temporal predicting typically take place on block or macroblock level, as described above.

[0026] With reference to Figure 1, a typical multimedia streaming system will be described, which is a preferred system for applying the procedure of the invention.

[0027] A multimedia data streaming system typically comprises one or more multimedia sources 100, such as a video camera and a microphone, or video image or computer graphic files stored in a memory carrier. Raw data obtained from the different multimedia sources 100 is combined into a multimedia file in an encoder 102, which can also be referred to as an editing unit. The raw data arriving from the one or more multimedia sources 100 is first captured using capturing means 104 included in the encoder 102, which capturing means can be typically implemented as different interface cards,

driver software, or application software controlling the function of a card. For example, video data may be captured using a video capture card and the associated software. The output of the capturing means 104 is typically either an uncompressed or slightly compressed data flow, for example uncompressed video frames of the YUV 4:2:0 format or motion-JPEG image format, when a video capture card is concerned.

[0028] An editor 106 links different media flows together to synchronize video and audio flows to be reproduced simultaneously as desired. The editor 106 may also edit each media flow, such as a video flow, by halving the frame rate or by reducing spatial resolution, for example. The separate, although synchronized, media flows are compressed in a compressor 108, where each media flow is separately compressed using a compressor suitable for the media flow. For example, video frames of the YUV 4:2:0 format may be compressed using the low bit rate video coding according to the ITU-T recommendation H.263 or H.26L. The separate, synchronized and compressed media flows are typically interleaved in a multiplexer 110, the output obtained from the encoder 102 being a single, uniform bit flow that comprises data of a plural number of media flows and that may be referred to as a multimedia file. It is to be noted that the forming of a multimedia file does not necessarily require the multiplexing of a plural number of media flows into a single file, but the streaming server may interleave the media flows just before transmitting them.

[0029] The multimedia files are transferred to a streaming server 112, which is thus capable of carrying out the streaming either as real-time streaming or in the form of progressive downloading. In progressive downloading the multimedia files are first stored in the memory of the server 112 from where they may be retrieved for transmission as need arises. In real-time streaming the editor 102 transmits a continuous media flow of multimedia files to the streaming server 112, and the server 112 forwards the flow directly to a client 114. As a further option, real-time streaming may also be carried out such that the multimedia files are stored in a storage that is accessible from the server 112, from where real-time streaming can be driven and a continuous media flow of multimedia files is started as need arises. In such case, the editor 102 does not necessarily control the streaming by any means. The streaming server 112 carries out traffic shaping of the multimedia data as regards the bandwidth available or the maximum decoding and playback rate

of the client 114, the streaming server being able to adjust the bit rate of the media flow for example by leaving out B-frames from the transmission or by adjusting the number of the scalability layers. Further, the streaming server 112 may modify the header fields of a multiplexed media flow to reduce their size and encapsulate the multimedia data into data packets that are suitable for transmission in the telecommunications network employed. The client 114 may typically adjust, at least to some extent, the operation of the server 112 by using a suitable control protocol. The client 114 is capable of controlling the server 112 at least in such a way that a desired multimedia file can be selected for transmission to the client, in addition to which the client is typically capable of stopping and interrupting the transmission of a multimedia file.

[0030] When the client 114 is receiving a multimedia file, the file is first supplied to a demultiplexer 116, which separates the media flows comprised by the multimedia file. The separate, compressed media flows are then supplied to a decompressor 118 where each separate media flow is decompressed by a decompressor suitable for each particular media flow. The decompressed and reconstructed media flows are supplied to a playback unit 120 where the media flows are rendered at a correct pace according to their synchronization data and supplied to presentation means 124. The actual presentation means 124 may comprise for example a computer or mobile station display, and loudspeaker means. The client 114 also typically comprises a control unit 122 that the end user can typically control through a user interface and that controls both the operation of the server, through the above-described control protocol, and the operation of the playback unit 120, on the basis of the instructions given by the end user.

[0031] It is to be noted that the transfer of multimedia files from the streaming server 112 to the client 114 takes place through a telecommunications network, the transfer path typically comprising a plural number of telecommunications network elements. It is therefore possible that there is at least some network element that can carry out traffic shaping of multimedia data with regard to the available bandwidth or the maximum decoding and playback rate of the client 114 at least partly in the same way as described above in connection with the streaming server.

[0032] Scalable coding will be described below with reference to a preferred embodiment of the invention and an example illustrated in Figure 2. Figure 2 shows part of a compressed video sequence having a first frame 200,

which is an INTRA frame, or I-frame, and thus an independently determined video frame the image information of which is determined without using motion-compensated temporal prediction. The I-frame 200 is placed on a first scalability layer, which may be referred to as an INTRA layer. Each scalability layer is assigned a unique identifier, such as the layer number. The INTRA layer may therefore be given the number 0, for example, or some other alphanumeric identifier, for example a letter, or a combination of a letter and a number.

[0033] Correspondingly, sub-sequences consisting of groups of one or more video frames are determined for each scalability layer, at least one of the images in a group (typically the first or the last one) being temporally predicted at least from a video frame of another, sub-sequence of typically either a higher or the same scalability layer, the rest of the video frames being temporally predicted either from only the video frames of the same sub-sequence, or possibly also from one or more video frames of said second sub-sequence. A sub-sequence may de decoded independently, irrespective of other sub-sequences, except said second sub-sequence. The sub-sequences of each scalability layer are assigned a unique identifier using for example consecutive numbering starting with number 0 given for the first sub-sequence of a scalability layer. Since the I-frame 200 is determined independently and can also be decoded independently upon reception, irrespective of other image frames, it also forms in a way a separate sub-sequence.

[0034] An essential aspect of the present invention is therefore to determine each sub-sequence in terms of those sub-sequences the sub-sequence is dependent on. In other words, a sub-sequence comprises information about all the sub-sequences that have been directly used for predicting the image frames of the sub-sequence in question. This information is signalled in a video sequence bit stream, preferably separate from the actual image information, and therefore the image data of the video sequence can be preferably adjusted because it is easy to determine video sequence portions that are to be independently decoded and can be removed without affecting the decoding of the rest of the image data.

[0035] Next, within each sub-sequence, the video frames of the sub-sequence are given image numbers, using for example consecutive numbering that starts with the number 0 given to the first video frame of the sub-sequence. Since the I-frame 200 also forms a separate sub-sequence, its

image number is 0. In Figure 2, the I-frame 200 shows the type (I), sub-sequence identifier and image number (0.0) of the frame.

[0036] Figure 2 further shows a next I-frame 202 of the INTRA layer, the frame thus being also an independently determined video frame that has been determined without using motion-compensated temporal prediction. The temporal transmission frequency of I-frames depends on many factors relating to video coding, image information contents and the bandwidth to be used, and, depending on the application or application environment, I-frames are transmitted in a video sequence at intervals of 0.5 to 10 seconds, for example. Since the I-frame 202 can be independently decoded, it also forms a separate sub-sequence. Since this is the second sub-sequence in the INTRA-layer, the consecutive numbering of the sub-sequence identifier of the I-frame 202 is 1. Further, since the I-frame 202 also forms a separate sub-sequence, i.e. it is the only video frame in the sub-sequence, its image number is 0. The I-frame 202 can thus be designated with identifier (I.1.0.) Correspondingly, the identifier of the next I-frame on the INTRA layer is (I.2.0.), etc. As a result, only independently determined I-frames in which the image information is not determined using motion-compensated temporal prediction are coded into the first scalability layer, i.e. the INTRA layer. The sub-sequences can also be determined using other kind of numbering or other identifiers, provided that the sub-sequences can be distinguished from one another.

[0037] The next scalability layer, which has a layer number 1, for example, and which may be referred to as the base layer, comprises coded, motion-compensated INTER or P-frames typically predicted only from previous image frames, i.e. in this case from the I-frames of an upper INTRA layer. The image information of the first P-frame 204 of the base layer shown in Figure 2 is determined using the I-frame 200 of the INTRA layer. A P-frame 204 begins the first sub-sequence of the base layer, and therefore the sub-sequence identifier of the P-frame 204 is 0. Further, since the P-frame 204 is the first image frame of the first sub-sequence of the base layer, the image number of the P-frame 204 is 0. The P-frame 204 can thus be identified with (P.0.0).

[0038] The temporally succeeding P-frame 206 of the base layer is predicted from the previous P-frame 204. The P-frames 204 and 206 thus belong to the same sub-sequence, whereby the P-frame 206 also receives the sub-sequence identifier 0. Since the P-frame 206 is the second image frame in

the sub-sequence 0, the image number of the P-frame 206 is 1, and the P-frame 206 can be identified with (P.0.1).

[0039] The scalability layer succeeding the base layer and having the layer number 2, is called enhancement layer 1. This layer comprises coded, motion-compensated P-frames predicted only from previous image frames, in this case either from I-frames of the INTRA layer or P-frames of the base layer. Figure 2 shows a first image frame 208 and a second image frame 210 of enhancement layer 1, which are both predicted only from the first image frame 200 of the INTRA layer. The P-frame 208 begins the first sub-sequence of enhancement layer 1, the sub-sequence identifier of the P-frame thus being 0. Further, since the P-frame 208 is the first and only image frame in said sub-sequence, the P-frame 208 receives the image number 0. The P-frame 208 can thus be identified with (P.0.0.).

[0040] Since the second image frame 210 is also predicted only from the first image frame 200 of the INTRA layer, the P-frame 210 begins the second sub-sequence of enhancement layer 1 and the sub-sequence identifier of the P-frame 210 is therefore 1. Since the P-frame 210 is the first image frame in the sub-sequence, the image number of the P-frame 210 is 0. The P-frame can thus be identified with (P.1.0.). The temporally succeeding P-frame 212 of enhancement layer 1 is predicted from the previous P-frame 210. The P-frames 210 and 212 thus belong to the same sub-sequence, and therefore the P-frame also receives the sub-sequence identifier 1. The P-frame 212 is the second image frame in the sub-sequence 1, and therefore the P-frame receives the image number 1 and can be identified with (P.1.1).

[0041] The temporally fourth image frame 214 of enhancement layer 1 is predicted from the first image frame 204 of the base layer. The P-frame 214 thus begins a third sub-sequence of enhancement layer 1 and therefore the P-frame 214 receives the sub-sequence identifier 2. Further, since the P-frame 214 is the first and only image frame in the sub-sequence, the image number of the P-frame 214 is 0. The P-frame 208 can therefore be identified with (P.2.0).

[0042] Also the temporally fifth image frame 216 of enhancement layer 1 is predicted only from the first image frame 204 of the base layer, the P-frame 216 thus beginning the fourth sub-sequence of enhancement layer 1, and the sub-sequence identifier of the P-frame 216 is 3. In addition, since the P-frame 216 is the first one in the sub-sequence in question, the image

number of the P-frame 216 is 0. The P-frame 216 can therefore be identified with (P.3.0). The temporally following P-frame 218 of enhancement layer 1 is predicted from the previous P-frame 216. The P-frames 216 and 218 thus belong to the same sub-sequence, and the sub-sequence identifier of the P-frame 218 is also 3. Since the P-frame 218 is the second image frame in the sub-sequence 3, the image number of the P-frame 218 is 1 and the identifier of the P-frame 218 is (P.3.1).

[0043] For simplicity and clarity of illustration the above disclosure only relates to I- and P-frames. However, a person skilled in the art will find it apparent that the scalable video coding of the invention can also be implemented using other known image frame types, such as the above described B-frames and at least SI-frames, SP-frames and MH-frames. SI-frames correspond to I-frames, but together with an SP-frame they allow an identical image to be reconstructed. An SP-frame, in turn, is a P-frame subjected to a particular coding that allows an identical image to be reconstructed together with an SI-frame or another SP-frame. SP-frames are typically placed into a video sequence into points where an access point or a scanning point is desired, or where the changing of the coding parameters of the video stream should be possible. The frames can also be used for error correction and for increasing error tolerance. SP-frames are otherwise similar to ordinary P-frames predicted from previous frames, except that they are defined so that they can be replaced by another video frame of the SP- or SI-type, the result of the decoding of the new frame being identical with the decoding result of the original SP-frame that was in the video stream. In other words, a new SP-frame that is used for replacing the one that was in the video stream is predicted from another sequence or video stream and yet the reconstructed frame has identical contents. SP-frames are described for example in the Applicant's earlier application PCT/FI02/00004.

[0044] Similarly as B-frames, macroblocks of MH (Multi Hypothesis) frames, based on motion-compensated prediction, are predicted from two other frames, which are not, however, necessarily located next to an MH-frame. More precisely, the predicted macroblocks are computed as an average of two macroblocks of two other frames. Instead of two frames, MH-frame macroblocks can naturally be also predicted from one other frame. Reference images may change according to macroblock, in other words, all macroblocks

in one and the same image are not necessarily predicted using the same frames.

[0045] A sub-sequence thus covers a specific period of time in a video sequence. The sub-sequences of the same layer or of different layers may be partly or entirely overlapping. If there are temporally overlapping image frames on the same layer, the frames are interpreted as alternative presentations of the same image content and therefore any mode of image presentation can be used. On the other hand, if there are temporally overlapping image frames on different layers, they form different presentations of the same image content, and therefore presentations differ in image quality, i.e. the quality of image is better on a lower layer.

[0046] The above disclosure referring to Figure 2 illustrates a scalable coding arrangement and a hierarchical structure and numbering of image frames according to a preferred embodiment of the invention. In this embodiment the INTRA-layer only comprises I-frames and the base layer can only be decoded using the information received from the INTRA-layer. Correspondingly, the decoding of enhancement layer 1 typically requires information from both the base layer and the INTRA-layer.

[0047] The number of the scalability layers is not restricted to three, as above, but any number of enhancement layers that is considered necessary for producing sufficient scalability may be used. Consequently, the layer number of enhancement layer 2 is four, that of enhancement layer 3 is five, etc. Since some of the image frames in the above example are given the same identifier (e.g. the identifier of both image frames 204 and 208 is (P.0.0)), by including the layer number in the identifier each image frame can be uniquely identified and, at the same time, the dependencies of each image frame on other image frames is preferably determined. Each image frame is thus uniquely identified, the identifier of image frame 204, for example, being (P.1.0.0), or simply (1.0.0) and, correspondingly, that of image 208 being (P.2.0.0), or (2.0.0).

[0048] According to a preferred embodiment of the invention, the number of a reference image frame is determined according to a specific, pre-determined alpha-numeric series, for example as an integer between 0 and 255. When the parameter value achieves the maximum value N (e.g. 255) in the series concerned, the determining of the parameter value starts from the beginning, i.e. from the minimum value of the series (e.g. 0). An image frame is

thus uniquely identified within a specific sub-sequence up to the point where the same image number is used again. The sub-sequence identifier can also be determined according to a specific, predetermined arithmetic series. When the value of the sub-sequence identifier achieves the maximum value N of the series, the determining of the identifier starts again from the beginning of the series. However, a sub-sequence cannot be assigned an identifier that is still in use (within the same layer). The series in use may also be determined in another way than arithmetically. One alternative is to assign random sub-sequence identifiers, taking into account that an assigned identifier is not be used again.

[0049] A problem in the numbering of image frames arises when the user wishes to start browsing a video file in the middle of a video sequence. Such situations occur for example when the user wishes to browse a locally stored video file backward or forward or to browse a streaming file at a particular point; when the user initiates the playback of a streaming file from a random point; or when a video file that is to be reproduced is detected to contain an error that interrupts the playback or requires the playback to be resumed from a point following the error. When the browsing of a video file is resumed from a random point after previous browsing, discontinuity typically occurs in the image numbering. The decoder typically interprets this as unintentional loss of image frames and unnecessarily tries to reconstruct the image frames suspected as lost.

[0050] According to a preferred embodiment of the invention, this can be avoided in the decoder by defining an initiation image in an independently decodable Group of Pictures GOP that is activated at the random point of the video file, and the number of the initiation image is set at zero. This independently decodable image group can thus be a sub-sequence of an INTRA-layer, for example, in which case an I-frame is used as the initiation image, or, if scaling originating from the base layer is employed, the independently decodable image group is a sub-sequence of the base layer, in which case the first image frame of the sub-sequence, typically an I-frame, is usually used as the initiation image. Consequently, when activated at a random point, the decoder preferably sets the identifier of the first image frame, preferably an I-frame, of the independently decodable sub-sequence at zero. Since the sub-sequence to be decoded may also comprise other image frames whose identifier is zero (for example when the above described alpha-

numeric series starts from the beginning), the beginning of the sub-sequence, i.e. its first image frame, can be indicated to the decoder for example by a separate flag added to the header field of a slice of the image frame. This allows the decoder to interpret the image numbers correctly and to find the correct image frame that initiates the sub-sequence from the video sequence image frames.

[0051] The above numbering system provides only one example of how the unique image frame identification of the invention can be carried out so that interdependencies between the image frames are indicated at the same time. However, video coding methods in which the method of the invention can be applied, such as video coding methods according to the ITU-T standards H.263 and H.26L, employ code tables, which in turn use variable length codes. When variable length codes are used for coding layer numbers, for example, a lower code word index, i.e. a smaller layer number, signifies a shorter code word. In practice the scalable coding of the invention will be used in most cases in such a way that the base layer will consist significantly more image frames than the INTRA-layer. This justifies the use of a lower index, i.e. a smaller layer number, on the base layer than on the INTRA-layer, because the amount of coded video data is thereby advantageously reduced. Consequently, the INTRA-layer is preferably assigned layer number 1 and the base layer is given layer number 0. Alternatively, the code can be formed by using fewer bits for coding the base layer number than the INTRA-layer number, in which case the actual layer number value is not relevant in view of the length of the code created.

[0052] Further, according to a second preferred embodiment of the invention, when the number of the scalability layers is to be kept low, the first scalability layer in particular can be coded to comprise both the INTRA-layer and the base layer. From the point of view of coding hierarchy, the simplest way to conceive this is to leave out the INTRA-layer altogether, and to provide the base layer with coded frames consisting of both independently defined I-frames, the image information of which has not been determined using motion-compensated temporal prediction, and image frames predicted from previous frames, which image frames in this case are motion-compensated P-frames predicted from the I-frames of the same layer. The layer number 0 can thus still be used for the base layer and, if enhancement layers are coded into the video

sequence, enhancement layer 1 is assigned layer number 1. This is illustrated in the following, with reference to Figures 3a and 3b.

[0053] Figure 3a shows a non-scalable video sequence structure, in which all image frames are placed on the same scalability layer, i.e. the base layer. The video sequence comprises a first image frame 300 which is an I-frame (I.0.0) and which thus initiates a first sub-sequence. The image frame 300 is used for predicting a second image frame 302 of the sub-sequence, i.e. a P-frame (P.0.1), which is then used for predicting a third image frame 304 of the sub-sequence, i.e. a P-frame (P.0.2), which is in turn used for predicting the next image frame 306, i.e. a P-frame (P.0.3). The video sequence is then provided with an I-frame (I.1.0) coded therein, i.e. an I-frame 308, which thus initiates a second sub-sequence in the video sequence. This kind of non-scalable coding can be used for example when the application employed does not allow scalable coding to be used, or there is no need for it. In a circuit-switched videophone application, for example, channel bandwidth remains constant and the video sequence is coded in real-time, and therefore there is typically no need for scalable coding.

[0054] Figure 3b, in turn, illustrates an example of how scalability can be added, when necessary, to a combined INTRA- and base layer. Here, too, the video sequence base layer comprises a first image frame 310 which is an I-frame (I.0.0) and which initiates a first sub-sequence of the base layer. The image frame 310 is used for predicting a second image frame 312 of the sub-sequence, i.e. a P-frame (P.0.1), which is then used for predicting a third image frame 314 of the sub-sequence, i.e. a P-frame (P.0.2). Enhancement layer 1, however, is also coded into this video sequence and it comprises a first sub-sequence, the first and only image frame 316 of which is a P-frame (P.0.0), which is predicted from the first image frame 310 of the base layer. The first image frame 318 of a second sub-sequence of the enhancement layer is, in turn, predicted from the second image frame 312 of the base layer, and therefore the identifier of this P-frame is (P.1.0). The next image frame 320 of the enhancement layer is again predicted from the previous image frame 318 of the same layer and, therefore, it belongs to the same sub-sequence, its identifier thus being (P.1.1).

[0055] In this embodiment of the invention the sub-sequences of the base layer can be decoded independently, although a base layer sub-sequence may be dependent on another base layer sub-sequence. The

decoding of the base layer sub-sequences requires information from the base layer and/or from the second sub-sequence of enhancement layer 1, the decoding of the sub-sequences of enhancement layer 2 requires information from enhancement layer 1 and/or from the second sub-sequence of enhancement layer 2, etc. According to an embodiment, I-frames are not restricted to the base layer alone, but lower enhancement layers may also comprise I-frames.

[0056] The basic idea behind the above embodiments is that a sub-sequence comprises information about all the sub-sequences it is dependent on, i.e. about all sub-sequences that have been used for predicting at least one of the image frames of the sub-sequence in question. However, according to an embodiment it is also possible that a sub-sequence comprises information about all sub-sequences that are dependent on the sub-sequence in question, in other words, about all the sub-sequences in which at least one image frame has been predicted using at least one image frame of the sub-sequence in question. Since in the latter case the dependencies are typically determined temporally forward, image frame buffers can be advantageously utilized in the coding in a manner to be described later.

[0057] In all the above embodiments the numbering of the image frames is sub-sequence-specific, i.e. a new sub-sequence always starts the numbering from the beginning. The identification of an individual image frame thus requires the layer number, sub-sequence identifier and image frame number to be determined. According to a preferred embodiment of the invention, the image frames can be independently numbered using consecutive numbering in which successive reference image frames in the coding order are indicated with numbers incremented by one. As regards layer numbers and sub-sequence identifiers, the above-described numbering procedure can be used. This allows each image frame to be uniquely identified, when necessary, without using the layer number and sub-sequence identifier.

[0058] This is illustrated with the example shown in Figure 4a in which the base layer comprises a temporally first I-frame 400 (I.0.0). This frame is used for predicting a first image frame 402 of enhancement layer 1, i.e. (P.0.1), which is then used for predicting a second image frame 404 belonging to the same sub-sequence (with sub-sequence identifier 0), i.e. (P.0.2), which is used for predicting a third image frame 406 of the same sub-

sequence, i.e. (P.0.3), which is used for predicting a fourth image frame 408 (P.0.4) and, finally, the fourth frame for predicting a fifth image frame 410 (P.0.5). The temporally next video sequence image frame 412 is located on the base layer, where it is in the same sub-sequence as the I-frame 400, although temporally it is only the seventh coded image frame, and therefore its identifier is (P.0.6). The seventh frame is then used for predicting a first image frame 414 of the second sub-sequence of enhancement layer 1, i.e. (P.1.7), which is then used for predicting a second image frame 416 belonging to the same sub-sequence (with sub-sequence identifier 1), i.e. (P.1.8), which in turn used for predicting a third image frame 418 (P.1.9), the third for predicting a fourth image frame 420 (P.1.10) and, finally, the fourth for predicting a fifth image frame 422 (P.1.11) of the same sub-sequence. Again, the temporally next video sequence image frame 424 is located on the base layer, where it is in the same sub-sequence as the I-frame 400 and the P-frame 412, although temporally it is only the thirteenth coded image frame and therefore its identifier is (P.0.12). For clarity of illustration, the above description of the embodiment does not comprise layer identifiers, but it is apparent that in order to implement scalability, also the layer identifier must be signalled together with the video sequence, typically as part of the image frame identifiers.

[0059] Figures 4b and 4c show alternative embodiments for grouping the image frames of the video sequence shown in Figure 4a. The image frames in Figure 4b are numbered according to sub-sequence, i.e. a new sub-sequence always starts the numbering from the beginning (from zero). Figure 4c, in turn, employs image frame numbering which corresponds otherwise to that used in Figure 4a, except that the P-frames of the base layer are replaced by SP-frame pairs to allow for identical reconstruction of image information.

[0060] As stated above, the procedure of the invention can also be implemented using B-frames. One example of this is illustrated in Figures 5a, 5b and 5c. Figure 5a shows a video sequence in the time domain, the sequence comprising P-frames P1, P4 and P7, with B-frames placed between them, the interdependencies of the B-frames with regard to temporal predicting being shown with arrows. Figure 5b shows a preferred grouping of video sequence image frames in which the interdependencies shown in Figure 5a are indicated. Figure 5b illustrates sub-sequence-specific image frame numbering in which a new sub-sequence always starts the numbering of the

image frames from zero. Figure 5c, in turn, illustrates image frame numbering, which is consecutive in the order of temporal prediction, wherein the following reference frame always receives the next image number as the previously encoded reference frame. The image frame (B1.8) (and (B2.10)) does not serve as a reference prediction frame to any other frame, therefore it does not affect the image frame numbering.

[0061] The above examples illustrate different alternatives of how scalability of video sequence coding can be adjusted by using the method of the invention. From the point of view of the terminal device reproducing the video sequence, the more scalability layers are available, or the more scalability layers it is capable of decoding, the better the image quality. In other words, increase in the amount of image information and in the bit rate used for transferring the information improves the temporal or spatial resolution, or the spatial quality of the image data. Correspondingly, a higher number of scalability layers also sets considerably higher demands on the processing capacity of the terminal device performing decoding.

[0062] In addition, the above examples illustrate the advantage gained by using sub-sequences. With image frame identifiers, the dependencies of each image frame from other image frames in the sub-sequence are indicated in an unambiguous manner. A sub-sequence thus forms an independent whole that can be left out of a video sequence, when necessary, without affecting the decoding of subsequent image frames of the video sequence. In that case, only the image frames of the sub-sequence in question and of those sub-sequences on the same and/or lower scalability layers dependent on it are not decoded.

[0063] The image frame identifier data transmitted together with the video sequence are preferably included in the video sequence header fields or in the header fields of the transfer protocol to be used for transferring the video sequence. In other words, the identifier data of predicted image frames are not included in the image data of the coded video sequence, but always into the header fields, whereby the dependencies of the image frames can be detected without decoding the images of the actual video sequence. The identifier data of the image frames can be stored for example in the buffer memory of the streaming server as the video sequence is being coded for transmission. In addition, the sub-sequences can be independently decoded on each scalability

layer, because the image frames of a sub-sequence are not dependent on other sub-sequences of the same scalability layer.

[0064] According to an embodiment of the invention, the image frames comprised by a sub-sequence may thus depend also on other sub-sequences of the same scalability layer. This dependency must then be signalled for example to the streaming server carrying out traffic shaping, because interdependent sub-sequences located on the same layer cannot be separately removed from a video sequence to be transmitted. A preferred way to carry out the signalling is to include it in the image frame identifiers to be transmitted, for example by listing the layer-sub-sequence pairs the sub-sequence in question depends on. This also provides a preferred way of indicating dependency from another sub-sequence of the same scalability layer.

[0065] The above examples illustrate a situation where image frames are temporally predicted from previous image frames. In some coding methods, however, the reference picture selection has been further extended to also include the predicting of the image information of image frames from temporally succeeding image frames. Reference picture selection offers most diversified means for creating different temporally scalable image frame structures and allows the error sensitivity of the video sequence to be reduced. One of the coding techniques based on reference picture selection is INTRA-frame postponement. The INTRA-frame is not placed into its temporally "correct" position in the video sequence, but its position is temporally postponed. The video sequence image frames that are between the "correct" position of the INTRA-frame and its actual position are predicted temporally backward from the INTRA-frame in question. This naturally requires that uncoded image frames be buffered for a sufficiently long period of time so that all image frames that are to be displayed can be coded and arranged into their order of presentation. INTRA-frame transfer and the associated determining of sub-sequences in accordance with the invention are illustrated in the following with reference to Figure 6.

[0066] Figure 6a shows a video sequence part in which the INTRA-frame comprises a single I-frame 600, which is temporally transferred to the position shown in Figure 6, although the "correct" position of the I-frame in the video sequence would have been at the first image frame. The video sequence image frames between the "correct" position and the real position 600 are thus

temporally predicted backward from the I-frame 600. This is illustrated by a sub-sequence coded into enhancement layer 1 and having a first temporally backward predicted image frame 602, which is a P-frame (P.0.0). This frame is used for temporally predicting a previous image frame 604, i.e. a P-frame (P.0.1), which is used in turn for predicting an image frame 606, i.e. a P-frame (P.0.2), and, finally, the frame 606 for predicting an image frame 608, i.e. a P-frame (P.0.3), which is at the position that would have been the "correct" position of the I-frame 600 in the video sequence. Correspondingly, the I-frame 600 on the base layer is used for temporally forward prediction of a sub-sequence comprising four P-frames 610, 612, 614 and 616, i.e. P-frames (P.0.0), (P.0.1), (P.0.2) and (P.0.3).

[0067] The fact that in this example backward predicted image frames are placed on a lower layer than forward predicted image layers indicates that for purposes of illustration, backward predicted image frames are in this coding example considered subjectively less valuable than forward predicted image frames. Naturally the sub-sequences could both be placed on the same layer, in which case they would be considered equal, or a backward predicted sub-sequence could be on the upper layer, in which case it would be considered subjectively more valuable.

[0068] Figures 6b and 6c show some alternatives for coding a video sequence according to Figure 6a. In Figure 6b both forward and backward predicted sub-sequences are placed on the base layer, the I-frame being only located on the INTRA-layer. The forward predicted sub-sequence on this layer is thus the second sub-sequence and its sub-sequence identifier is 1. In Figure 6c, in turn, an I-frame and a forward predicted sub-sequence based on it are located on the base layer, while a backward predicted sub-sequence is located on enhancement layer 1.

[0069] Moreover, according to a preferred embodiment of the invention, the above-described scalability can be utilized for coding what is known as a scene transition into a video sequence. Video material, such as news reports, music videos and movie trailers, often comprise rapid cuts between separate image material scenes. Sometimes the cuts are abrupt, but often a procedure known as scene transition is used in which transfer from one scene to another takes place by dimming, wiping, mosaic dissolving or scrolling the image frames of a previous scene, and, correspondingly, by presenting those of a later scene. From the point of view of coding efficiency,

the video coding of a scene transition is often most problematic, because the image frames appearing during the scene transition comprise information on the image frames of both the terminating and the initiating scene.

[0070] A typical scene transition, fading, is carried out by gradually reducing the intensity or luminance of the image frames of a first scene to zero, while gradually increasing the intensity of the image frames of a second scene to its maximum value. This scene transition is referred to as cross-faded scene transition.

[0071] Generally speaking, a computer-made image can be thought of as consisting of layers, or image objects. Each object can be defined with reference to at least three information types: the structure of the image object, its shape and transparency, and the layering order (depth) in relation to the background of the image and to other image objects. Shape and transparency are often determined using what is known as an alpha plane, which measures opacity and the value of which is usually determined separately for each image object, possibly excluding the background, which is usually determined as non-transparent. The alpha plane value of a non-transparent image object, such as the background, can thus be set at 1.0, whereas the alpha plane value of a fully transparent image object is 0.0. The values in between define the intensity of the visibility of a specific image object in a picture in proportion to the background and to other, at least partly overlapping, image objects that have a higher depth value than the image object in question.

[0072] The superimposition of image objects in layers according to their shape, transparency and depth position is referred to as scene composition. In practice the procedure is based on the use of weighted averages. First, the image object that is closest to the background, i.e. deepest according to its depth position, is placed onto the background and a combined image is formed of the two. The pixel values of the combined image are formed as an average weighted by the alpha plane values of the background image and the image object in question. The alpha plane value of the combined image is then set at 1.0, after which it serves as a background image for the next image object. The process continues until all image objects are attached to the image.

[0073] In the following, a procedure according to a preferred embodiment of the invention will be described in which video sequence scalability layers are combined with the above described image objects of

image frames and their information types to provide a scene transition with scalable video coding that also has good compression efficiency.

[0074] This embodiment of the invention is illustrated in the following by way of example and in a simplified manner by using cross-faded scene transition, on one hand, and abrupt scene transition, on the other hand, as examples. The image frames to be displayed during a scene transition are typically formed of two superimposed image frames, a first image frame comprising a first image scene and a second image frame a second scene. One of the image frames serves as the background image and other, which is referred to as a foreground image, is placed on top of the background image. The opacity of the background image, i.e. its non-transparency value, is constant. In other words, its pixel-specific alpha plane values are not adjusted.

[0075] In this embodiment of the invention, the background and foreground images are both defined according to scalability layer. This is illustrated in Figure 7, which shows an example of how image frames of two different scenes can be placed on scalability layers during a scene transition of the invention. Figure 7 shows a first image frame 700 of a first (terminating) scene positioned on the base layer. The image frame 700 may be either an I-frame containing image information that has not been determined using motion-compensated temporal predicting, or it may be a P-frame that is a motion-compensated image frame predicted from previous image frames. The coding of a second (initiating) scene starts during the temporally following image frame and, according to the invention, the image frames of the scene are also placed on the base layer. Remaining image frames 702, 704 of the second (terminating) scene are then placed on enhancement layer 1. These image frames are typically P-frames.

[0076] In this embodiment, the image frames of the second (initiating) scene are thus placed on the base layer, at least for the duration of the scene transition. The first image frame 706 of the scene is typically an I-frame, and it is used for temporally predicting the succeeding image frames of the scene. Consequently, the succeeding image frames of the second scene are temporally predicted frames, typically P-frames, such as frames 708 and 710 shown in Figure 7.

[0077] According to a preferred embodiment of the invention, this placing of image frames on scalability layers can be used for implementing a cross-faded scene transition by determining the image layer that is on the base

layer always as a background image of maximum opacity (100%), or non-transparency value. During a scene transition, image frames located on enhancement layers are placed onto the background image and their opacity is adjusted for example by means of suitable filters such that the frames gradually change from non-transparent to transparent.

[0078] In the video sequence of Figure 7, there are no image frames on the lower scalability layers during the first base layer image frame 700. For this time instant, the first image frame 700 is only coded into the video sequence.

[0079] The next image frame 706 of the base layer initiates a new (second) scene, during which the image frame 706 is provided with depth positioning that places it as the background image, and its opacity value is set to the maximum. Temporally simultaneously with the image frame 706 of the base layer, there is an image frame 702 of a terminating (first) scene on enhancement layer 1. To allow a cross-faded scene transition to be produced, the transparency of the frame 702 must be increased. The example of Figure 7 assumes that the opacity of the image frame 702 is set at 67% and, in addition, the image frame 702 is provided with depth positioning that determines it as a foreground image. For this time instant, an image combining the image frames 706 and 702 is coded into the video sequence, image 706 being visible as a weaker image on the background and image 702 as a stronger image at the front, because its opacity value is essentially high (67%).

[0080] During the temporally following image frame, there is a second image frame 708 of the second scene on the base layer, the frame 708 being thus correspondingly provided with depth positioning determining it as a background image, and its opacity value is set to the maximum. Enhancement layer 1 further comprises the last image frame 704 of a temporally simultaneously terminating (first) scene, the opacity value of the frame being set at 33% and, in addition, the image frame 704 being provided with depth positioning that determines it as a foreground image as well. Consequently, for this time instant, an image combined of the image frames 708 and 704 is coded into the video sequence, the image 708 being displayed as a stronger image on the background and the image 704 as a weaker image on the foreground, because the opacity value of the image 704 is no longer more than 33%.

[0081] During the temporally following image frame, the base layer comprises a third image frame 710 of the second scene. Since the first scene has terminated, only the image frame 710 is coded into the video sequence, and the displaying of the second scene continues from the frame 710.

[0082] The above disclosure describes, by way of example, the positioning of image frames according to the invention on scalability layers to implement cross-faded scene transition in a manner that is advantageous from the point of view of coding efficiency. However, it is possible that when a video sequence is being transmitted or decoded, a situation arises in which the bit rate of the video sequence must be adjusted according to the maximum value of the bandwidth and/or terminal device decoding rate available for data transfer. This kind of bit rate control causes problems when the scene transition is to be implemented using prior art video coding methods.

[0083] A preferred embodiment of the present invention now allows one or more scalability layers, or independently decodable sub-sequences included in them, to be removed from a video sequence, whereby the bit rate of the video sequence can be decreased and yet, at the same time, the video sequence can be decoded without reducing image frequency. In the image frame positioning according to Figure 7, this can be implemented by removing enhancement layer 1 from the video sequence. The video sequence is thus only used for displaying the image frames 700, 706, 708 and 710 of the base layer. In other words, a direct transition from the first (terminating) scene to the second (initiating) scene takes place in the form of an abrupt scene transition, i.e. directly from the image frame 700 of the first scene into the I-image frame 706 that initiates the second scene. The transition is thus not a cross-faded scene transition but an abrupt scene transition. Nevertheless, the scene transition can be carried out in an advantageous manner without affecting the quality of the video sequence image, and the viewer usually does not experience an abrupt scene transition carried out instead of a cross-faded scene transition in any way disturbing or faulty. On the contrary, since the prior art implementation does not allow scalability layers to be removed, scene transition would often require image frequency to be reduced, which the viewer would find jerky and disturbing.

[0084] The invention thus provides a preferred means for carrying out multimedia data traffic shaping in a streaming server comprising information about the different sub-sequences of a video sequence: their

average bit rate, location in relation to the entire video sequence, duration and their interdependencies regarding the layers. The streaming server also determines the maximum value of the bandwidth available for the data transfer and/or the decoding rate of the terminal device. On the basis of this information, the streaming server decides how many scalability layers and which sub-sequences are transmitted in the video sequence. Bit rate control can thus be carried out, when necessary, by making first a rough adjustment of the number of the scalability layers, after which finer sub-sequence-specific adjustment can be easily carried out. At its simplest, bit rate control means making sub-sequence-specific decisions on whether a particular sub-sequence will be added to a video sequence or removed from it. In case of removal it is advisable to remove entire sub-sequences from a video sequence, because the removal of separate images may cause errors in other images of the same sub-sequence. For the same reason, all sub-sequences of a lower enhancement layer should be left out if they are dependent on the removed sub-sequence of a higher layer. If there are interdependent sub-sequences on one and the same scalability layer, sub-sequences dependent on an earlier sub-sequence must be removed if the earlier sub-sequence is removed.

[0085] If the image frame identifier data are added to a video sequence that is to be transmitted, traffic shaping can also be carried out in a telecommunications network element to be used for the transfer of the video sequence, for example in an Internet router, in different gateways, or at a base station or base station controller of a mobile communications network. For the network element to be able to maintain and process the sub-sequence information, it must have extra memory and processing capacity. For this reason traffic shaping that is to be carried out in the network is perhaps most probably executed using simple processing methods, such as the DiffServ, i.e. differentiated services, procedure that is supported by some IP-based networks. In the DiffServ method, each IP data packet is assigned a priority, whereby data packets of a higher priority are relayed faster and more reliably to the recipient than packets of a lower priority. This is advantageously applied to the scalability of the invention by determining not only scalability-layer-specific, but also sub-sequence-specific priorities, which enables a highly advanced priorisation.

[0086] There are many alternatives for adding image frame identifier data to a video sequence that is to be transmitted. In addition, it is also

possible not to include any identifier data into the video sequence, in which case traffic shaping is only carried out at the streaming server. The identifier data can be included in the header fields of a video sequence, or in the header fields of the transfer protocol to be used, such as RTP (Real Time Protocol). According to a preferred embodiment, the identifier data can be transferred using a Supplemental Enhancement Information (SEI) mechanism. SEI provides a data delivery mechanism that is transferred synchronously with the video data content, thus assisting in the decoding and displaying of the video sequence. The SEI mechanism, particularly when used for transferring layer and sub-sequence information, is disclosed more in detail in the ITU-T standard document ITU-T Rec. H.264 (ISO/IEC 14496 – 10:2002), Annex D. In the cases, wherein a separate transfer protocol or mechanism is used for identifier data transfer, traffic shaping can be carried out also at one of the network elements of the transfer path. In addition, the receiving terminal device can control the decoding.

[0087] If the encoder or decoder supports reference picture selection, video sequence coding requires that decoded image frames be buffered before the coding so as to allow the relationships between different image frames to be temporally predicted from one or more other image frames. Image frame buffering can be arranged at least in two different ways, either as sliding windowing or as adaptive buffer memory control. In sliding windowing, M image frames that were coded last are used as a buffer. The frames in the buffer memory are in a decoded and reconstructed form, which allows them to be used as reference images in the coding. As the coding proceeds, the image frame buffering functions on the basis of the FIFO principle (First-In-First-Out). Images that are not used as reference images, such as conventional B-images, do not need to be stored in the buffer. Alternatively, the buffering can be also be implemented as adaptive buffer memory control, in which case the image buffering is not restricted to the FIFO principle, but image frames that are not needed can be removed from the buffer in the middle of the process, or, correspondingly, some image frames can be stored in the buffer for a longer period of time, if they are needed as reference images for later image frames. A known reference picture selection is implemented by indexing image frames that are in the buffer memory into a specific order, the image indices being then used to refer to an image in connection with motion-compensation, for example. This indexing method generally provides better compression

efficiency compared to using image numbers, for example, for referring to a specific image when motion-compensation reference images are to be signalled.

[0088] The above reference image indexing method is sensitive to transfer errors, because the buffers of the sender's encoder and the recipient's decoder must contain mutually corresponding reconstructed images in identical order to ensure that the encoder and decoder both form the same indexing order. If the image frames are indexed in different order in the buffers of the encoder and the decoder, an incorrect reference image may be used in the decoder. To prevent this, it is essential that the decoder can be controlled to take into account image frames and sub-sequences that the encoder has intentionally removed from the video sequence. In that case the image frame numbering may comprise gaps, which the decoder typically interprets as errors and tries to reconstruct the image frames interpreted as lost. For this reason, it is essential that the encoder is capable to inform the decoder that the discontinuities in the image numbering of the transmitted image frames are intentional.

[0089] In response to this, and provided that sliding windowing is used for buffering the image frames, the decoder enters into the buffer memory a number of image frames, the contents of which may be fully random, corresponding to the missing image numbers. These random image frames are then designated by an identifier "invalid" to indicate that the frames in question do not belong to the actual video sequence, but are only filler frames entered for purposes of buffer memory management. A filler frame can naturally be implemented using only memory indicators, i.e. no data is preferably entered into the buffer memory, but memory management is used merely to store a reference to a generic "invalid" frame. The entering of the image frames of the actual video sequence continues from the correct image frame number after the number of filler frames indicated by the missing image numbers has been entered into the buffer, which allows the buffer memories of the encoder and the decoder to be kept preferably in synchronism. If during decoding a reference to an image number is detected which is then found to indicate a filler frame located in the buffer, error correction actions are initiated in the decoder to reconstruct the actual reference image, for example by asking the encoder to re-transmit the reference image in question.

[0090] Further, the procedure of the invention allows separate buffer memories to be used on the different scalability layers, or, correspondingly, sub-sequence-specifically. Each scalability layer may thus have a separate buffer memory that is conceptually separate and functions on the basis of the sliding window principle. Similarly, each sub-sequence may also be provided with a conceptually separate buffer memory that also functions on the basis of the sliding window principle. This means that the buffer memory is always emptied when a sub-sequence terminates. Separate buffer memories can be used in a preferred manner for reducing the need for signalling in certain situations in which ordinary sliding window buffering would be inadequate and actively adaptive buffer memory management would need to be used instead.

[0091] The H.26L standard defines a picture order count as a picture position in output order. The decoding process specified in the H.26L standard uses picture order counts to determine default index orderings for reference pictures in B slices, to represent picture order differences between frames and fields for vector scaling in motion vector prediction and for implicit mode weighted prediction in B slices, and to determine when successive slices in decoding order belong to different pictures. The picture order count is coded and transmitted for each picture.

[0092] In one embodiment of the invention, the decoder uses the picture order count to conclude that pictures are temporally overlapping, i.e., pictures that have an equal picture order count are temporally overlapping. Preferably, the decoder outputs only the picture on the highest received layer. In the absence of layer information, the decoder concludes that the latest temporally overlapping picture in decoding order resides on highest received layer.

[0093] The above disclosure describes a procedure for coding video frames for the purpose of producing a scalable, compressed video sequence. The actual procedure is carried out in a video encoder, such as the compressor 108 of Figure 1, which may be any known video encoder. For example a video encoder according to the ITU-T recommendation H.263 or H.26L may be used, the video encoder being arranged to form, in accordance with the invention, a first sub-sequence into a video sequence, at least part of the sub-sequence being formed by coding I-frames; to form at least a second sub-sequence into the video sequence, at least part of the sub-sequence being formed by coding at least P- or B-frames, and at least one video frame of

the second sub-sequence being predicted from at least one video frame of the first sub-sequence; and to determine into the video sequence the identification data of at least the video frames of the second sub-sequence.

[0094] According to the procedure of the invention, each sub-sequence of a particular scalability layer is preferably independently decodable, naturally taking into account dependencies from higher scalability layers and possibly other sub-sequences of the same scalability layer. A scalably compressed video sequence such as the one described above can thus be decoded by decoding a first sub-sequence of a video sequence, at least part of the sub-sequence having been formed by coding at least I-frames, and by decoding at least a second sub-sequence of the video sequence, at least part of the second sub-sequence having been formed by coding at least P- or B-frames, and at least one video frame of the second sub-sequence having been predicted from at least one video frame of the first sub-sequence, and by determining the identification and dependency data of at least the video frames comprised by the second sub-sequence of the video sequence, and by reconstructing at least part of the video sequence on the basis of the sub-sequence dependencies.

[0095] The actual decoding takes places in the video decoder, such as the decompressor 118 of Figure 1, which may be any known video decoder. For example, a low bit rate video decoder according to the ITU-T recommendation H.263 or H.26L may be used, which in this invention is arranged to decode a first sub-sequence of a video sequence, at least part of the sub-sequence having been formed by coding I-frames; to decode at least a second sub-sequence of the video sequence, at least part of the second sub-sequence having been formed by coding at least P- or B-frames and at least one video frame of the second sub-sequence having been predicted from at least one video frame of the first sub-sequence. The video decoder is arranged to determine the identification and dependency data of at least the video frames comprised by the second sub-sequence of the video sequence and to reconstruct at least part of the video sequence on the basis of the dependencies of the sub-sequences.

[0096] An essential aspect in the operation of the streaming system of the invention is that the encoder and decoder are positioned at least so that the encoder is operationally connected to the streaming server and the decoder is operationally connected to the receiving terminal device. However,

the different elements of the streaming system, terminal devices in particular, may include functionalities that allow two-way transfer of multimedia files, i.e. transmission and reception. The encoder and decoder can thus be implemented in the form of what it known as a video codec integrating both encoder and decoder functionalities.

[0097] It is to be noted that according to the invention the functional elements of the above described streaming system and its elements, such as the streaming server, video encoder, video decoder and terminal are preferably implemented by means of software, by hardware solutions, or as a combination of the two. The coding and decoding methods of the invention are particularly suitable for implementation as computer software comprising computer-readable commands for executing the process steps of the invention. A preferred way of implementing the encoder and the decoder is to store them in a storage means as a program code that can be executed by a computer-like device, for example a personal computer (PC) or a mobile station, to provide coding/decoding functionalities on the device in question.

[0098] Another alternative is to implement the invention as a video signal comprising a scalably compressed video sequence which in turn comprises video frames coded according to at least a first and a second frame format, the video frames according to the first frame format being independent of other video frames, and the video frames of the second frame format being predicted from at least one of the other video frames. According to the invention, the video signal in question comprises at least a first sub-sequence, at least part of which has been formed by coding at least video frames of the first frame format; at least a second sub-sequence, at least part of which has been formed by coding at least video frames of the second frame format; and at least one video frame of the second sub-sequence having been predicted from at least one video frame of the first sub-sequence; and at least one data field that determines video frames belonging to the second sub-sequence.

[0099] It is apparent to a person skilled in the art that as technology advances the basic idea of the invention can be implemented in various ways. The invention and its embodiments are therefore not restricted to the above examples, but they may vary within the scope of the claims.

CLAIMS

1. A method for encoding a video sequence comprising an independent sequence of image frames, wherein at least one reference image frame is predictable from at least one previous image frame that is earlier than the previous reference image frame in decoding order, the method comprising:

encoding into the video sequence an indication of at least one image frame, which is the first image frame, in decoding order, of the independent sequence, said at least one reference image frame being included in the sequence.

2. A method according to claim 1, further comprising:

encoding the indication into the video sequence as a separate flag included in the header of a slice.

3. A method according to claim 1, further comprising:

encoding identifier values for image frames according to a numbering scheme, and

resetting the identifier value for the indicated first image frame of an independent sequence.

4. A method according to claim 1, further comprising:

encoding into the video sequence an identifier value for the independent sequence.

5. A video encoder for encoding a video sequence comprising an independent sequence of image frames, wherein at least one reference image frame is predictable from at least one previous image frame that is earlier than the previous reference image frame in decoding order, the video encoder being arranged to

encode into the video sequence an indication of at least one image frame, which is the first image frame, in decoding order, of the independent sequence, said at least one reference image frame being included in the sequence.

6. A video encoder according to claim 5, wherein the video encoder is arranged to

encode the indication into the video sequence as a separate flag included in the header of a slice.

7. A video encoder according to claim 5, wherein the video encoder is arranged to

encode identifier values for image frames according to a numbering scheme, and

reset the identifier value for the indicated first image frame of an independent sequence.

8. A video encoder according to claim 5, wherein the video encoder is arranged to

encode into the video sequence an identifier value for the independent sequence.

9. A computer program product, stored on a computer readable medium and executable in a data processing device, for encoding a video sequence comprising an independent sequence of image frames, wherein at least one reference image frame is predictable from at least one previous image frame that is earlier than the previous reference image frame in decoding order, the computer program product comprising:

a computer program code for encoding into the video sequence an indication of at least one image frame, which is the first image frame, in decoding order, of the independent sequence, said at least one reference image frame being included in the sequence.

10. A method for decoding a compressed video sequence, wherein at least one reference image frame is predictable from at least one previous image frame that is earlier than the previous reference image frame in decoding order, the method comprising

decoding from the video sequence an indication of at least one image frame, which is the first image frame, in decoding order, of an independent sequence, said sequence including the at least one reference image frame; and

starting the decoding of the video sequence from said first image frame of the independent sequence, whereby the video sequence is decoded without prediction from any image frame decoded prior to said first image frame.

11. A method according to claim 10, wherein the indication is a separate flag included in the header of a slice.

12. A method according to claim 10, further comprising:

decoding identifier values for image frames according to a numbering scheme, and

resetting the identifier value for the indicated first image frame of an independent sequence.

13. A video decoder for decoding a compressed video sequence, wherein at least one reference image frame is predictable from at least one previous image frame that is earlier than the previous reference image frame in decoding order, the video decoder being arranged to

decode from the video sequence an indication of at least one image frame, which is the first image frame, in decoding order, of an independent sequence, said sequence including the at least one reference image frame; and

start the decoding of the video sequence from said first image frame of the independent sequence, whereby the video sequence is decoded without prediction from any image frame decoded prior to said first image frame.

14. A computer program product, stored on a computer readable medium and executable in a data processing device, for decoding a compressed video sequence, wherein at least one reference image frame is predictable from at least one previous image frame that is earlier than the previous reference image frame in decoding order, the computer program product comprising:

a computer program code for decoding from the video sequence an indication of at least one image frame, which is the first image frame, in decoding order, of an independent sequence, said sequence including the at least one reference image frame; and

a computer program code for starting the decoding of the video sequence from said first image frame of the independent sequence, whereby the video sequence is decoded without prediction from any image frame decoded prior to said first image frame.

15. A video signal comprising a compressed video sequence including an independent sequence of image frames, wherein at least one reference image frame is predictable from at least one previous image frame that is earlier than the previous reference image frame in decoding order, the video signal further comprising:

an indication of at least one image frame, which is the first image frame, in decoding order, of the independent sequence, said at least one reference image frame being included in the sequence.

16. A video signal according to claim 15, the video signal further comprising an identifier value for the independent sequence.

ABSTRACT

A method for encoding a video sequence comprising an independent sequence of image frames, wherein at least one reference image frame is predictable from at least one previous image frame that is earlier than the previous reference image frame in decoding order. An indication of at least one image frame is encoded into the video sequence, which indicated image frame is the first image frame, in decoding order, of the independent sequence, said at least one reference image frame being included in the sequence. In the decoding phase, the indication of at least one image frame is decoded from the video sequence, and the decoding of the video sequence is started from said first image frame of the independent sequence, whereby the video sequence is decoded without prediction from any image frame decoded prior to said first image frame.

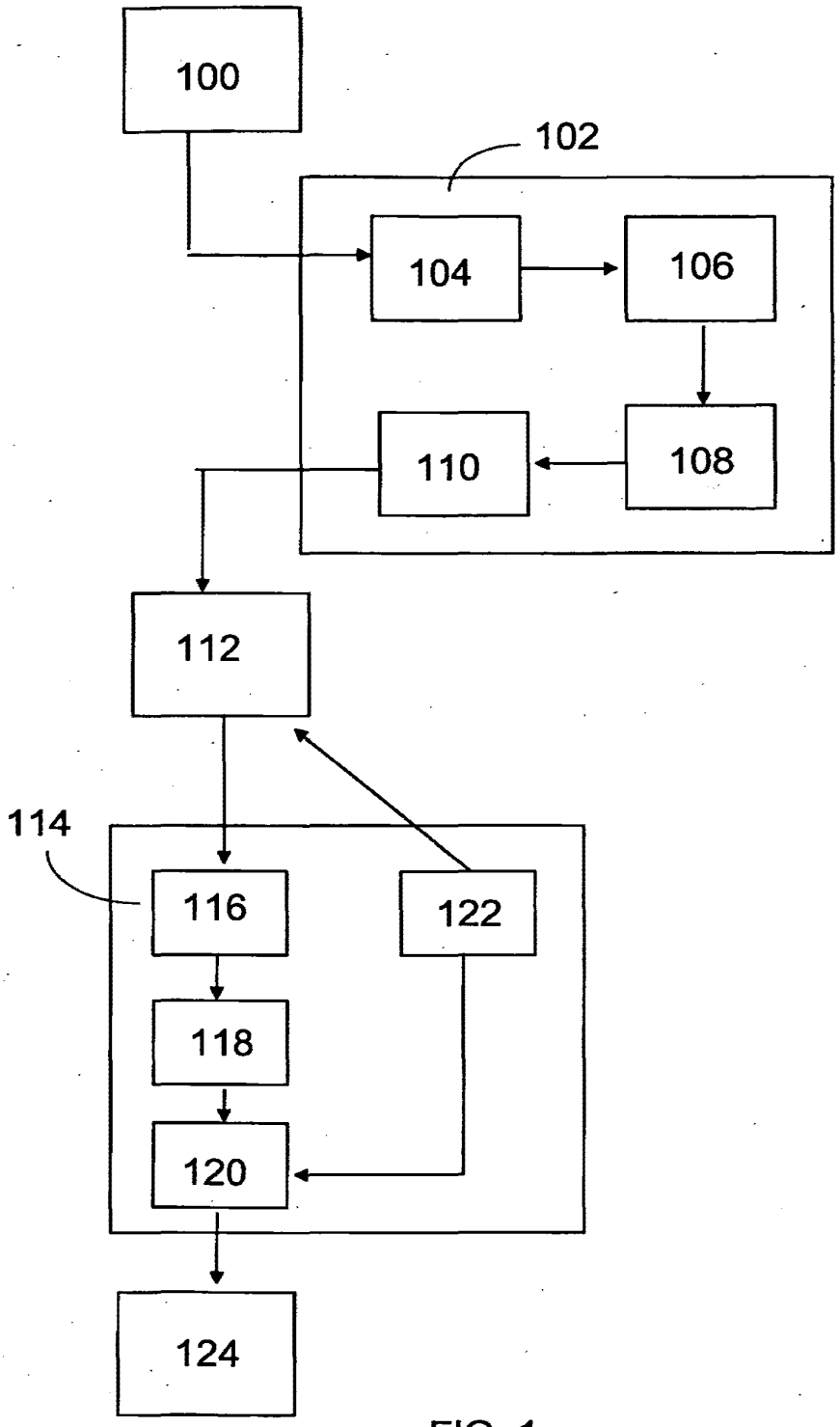


FIG. 1

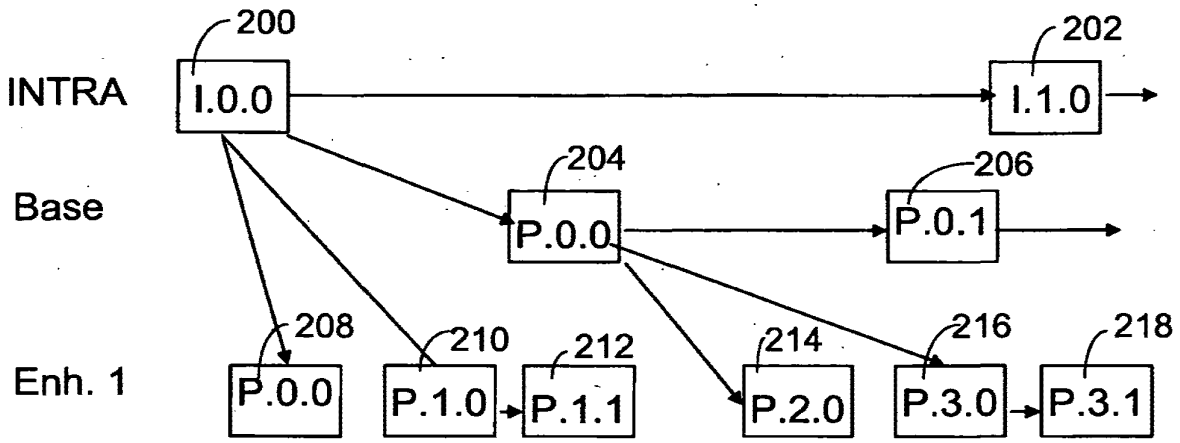


FIG. 2

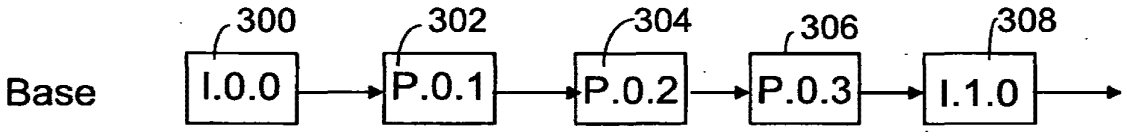


FIG. 3a

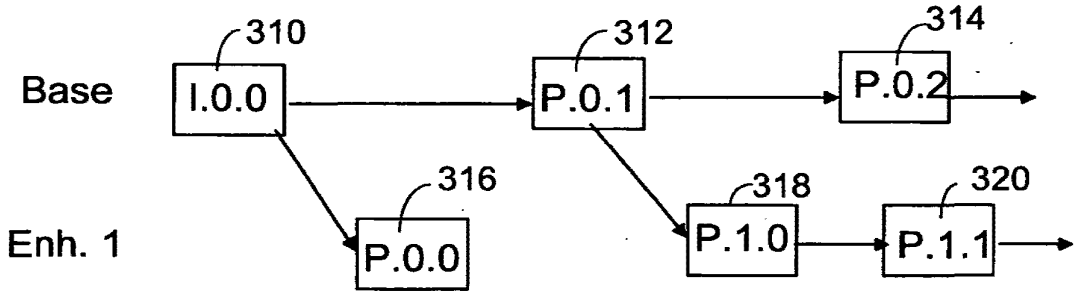


FIG. 3b

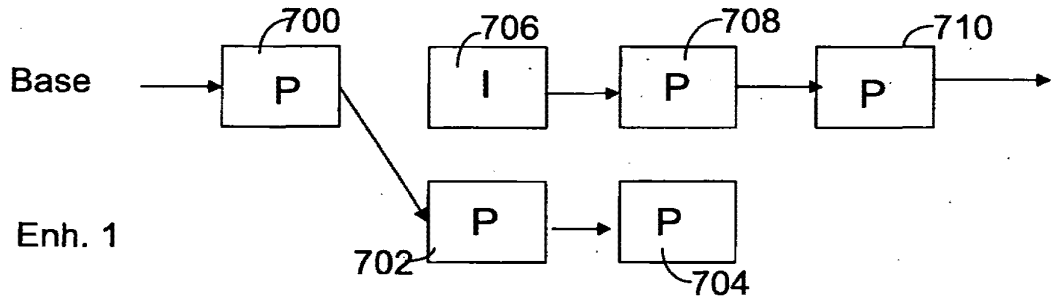


FIG. 7

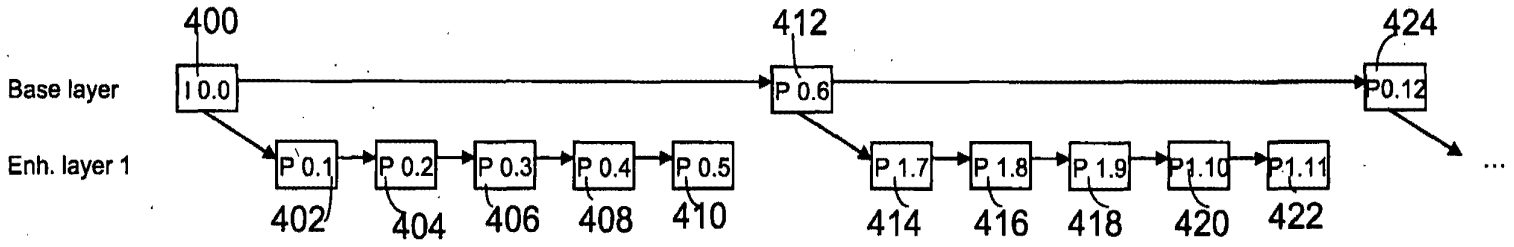
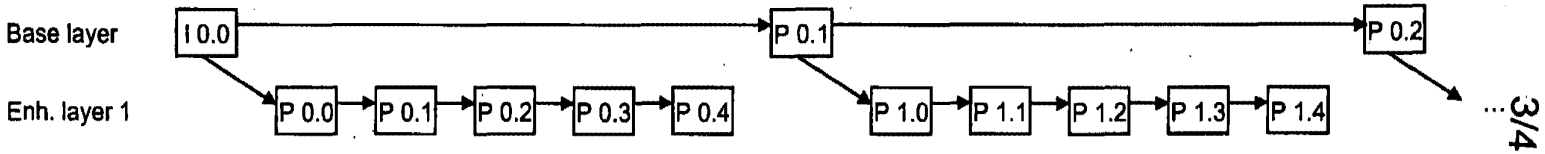


FIG. 4a



3/4

FIG. 4b

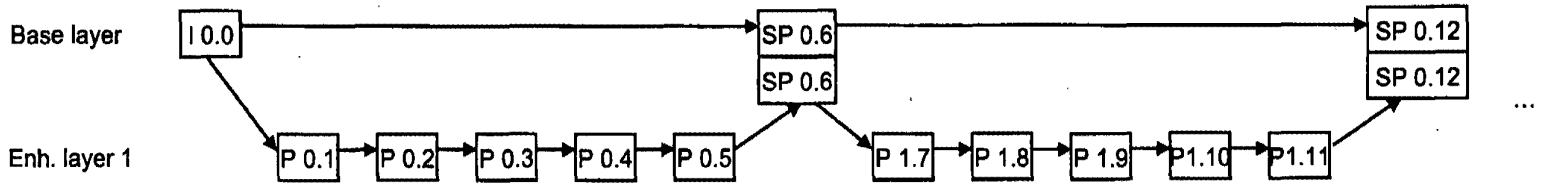


FIG. 4c

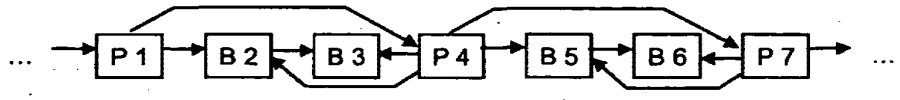


FIG. 5a

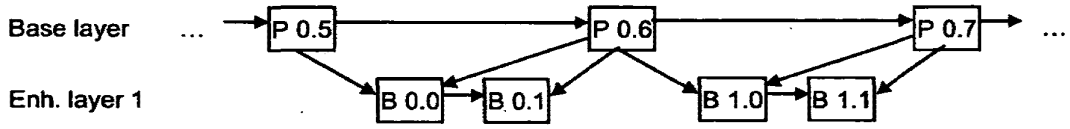


FIG. 5b

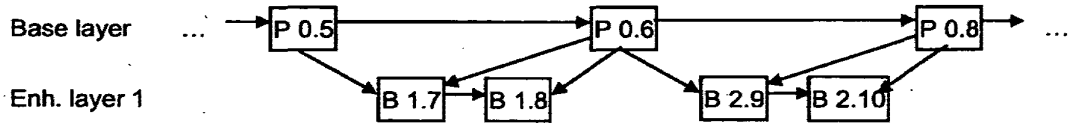


FIG. 5c

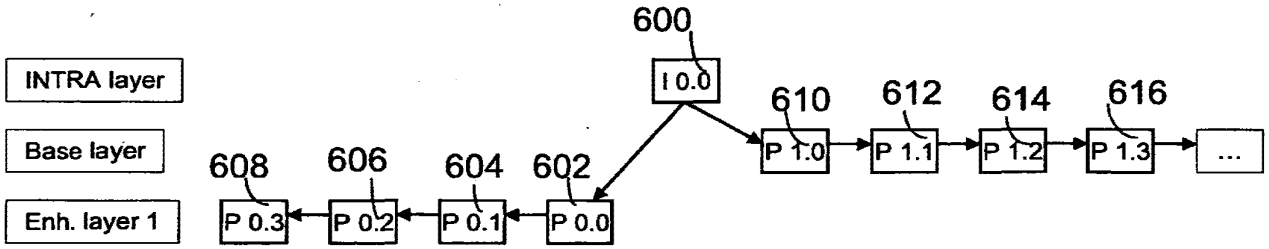


FIG. 6a

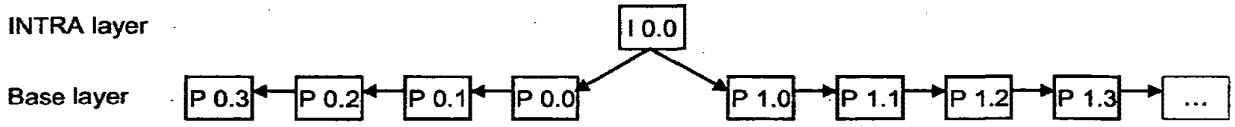


FIG. 6b

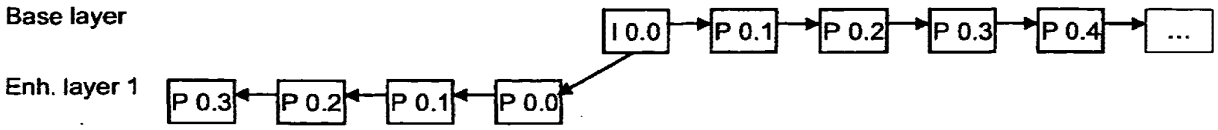


FIG. 6c

CRAWFORD MAUNU PLLC

United States Patent Application

COMBINED DECLARATION AND POWER OF ATTORNEY

As a below named inventor I hereby declare that: my residence, post office address and citizenship are as stated below next to my name; that

I verily believe I am the original; first and sole inventor (if only one name is listed below) or a joint inventor (if plural inventors are named below) of the subject matter which is claimed and for which a patent is sought on the invention entitled: **GROUPING OF IMAGE FRAMES IN VIDEO CODING.**

The specification of which
a. ☐ is attached hereto
b. ☒ is entitled **GROUPING OF IMAGE FRAMES IN VIDEO CODING,** having attorney docket number {**KOLS.021PA** (NC 35273US)}.
c. ☒ was filed on November 29, 2002 as application serial no. 10/306,942 and was amended on (if applicable) (in the case of a PCT-filed application) described and claimed in international no. filed and as amended on (if any), which I have reviewed and for which I solicit a United States patent.

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to the patentability of this application in accordance with Title 37, Code of Federal Regulations, § 1.56 (attached hereto).

I hereby claim foreign priority benefits under Title 35, United States Code, § 119/365 of any foreign application(s) for patent or inventor's certificate listed below and have also identified below any foreign application for patent or inventor's certificate having a filing date before that of the application on the basis of which priority is claimed:

a. ☐ no such applications have been filed.
b. ☒ such applications have been filed as follows:

FOREIGN APPLICATION(S), IF ANY, CLAIMING PRIORITY UNDER 35 USC § 119			
COUNTRY	APPLICATION NUMBER	DATE OF FILING (day, month, year)	DATE OF ISSUE (day, month, year)
FINLAND	20020127	23 JANUARY 2002	
ALL FOREIGN APPLICATION(S), IF ANY, FILED BEFORE THE PRIORITY APPLICATION(S)			
COUNTRY	APPLICATION NUMBER	DATE OF FILING (day, month, year)	DATE OF ISSUE (day, month, year)

I hereby claim the benefit under Title 35, United States Code, § 120/365 of any United States and PCT international application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code, § 112, I acknowledge the duty to disclose material information as defined in Title 37, Code of Federal Regulations, § 1.56(a) which occurred between the filing date of the prior application and the national or PCT international filing date of this application.

U.S. APPLICATION NUMBER	DATE OF FILING (day, month, year)	STATUS (patented, pending, abandoned)

I hereby claim the benefit under Title 35, United States Code § 119(e) of any United States provisional application(s) listed below:

U.S. PROVISIONAL APPLICATION NUMBER	DATE OF FILING (Day, Month, Year)

I hereby appoint the following attorney(s) and/or patent agent(s) to prosecute this application and to transact all business in the Patent and Trademark Office connected herewith:

Funk, Steven R.	Reg. No. 37,830	Crawford, Robert J.	Reg. No. 32,122	Maunu, LeRoy D.	Reg. No. 35,274
Hollingsworth, Mark A.	Reg. No. 38,491	Curtin, Eric J.	Reg. No. 47,511	Davis, Clara	Reg. No. 50,495

I hereby authorize them to act and rely on instructions from and communicate directly with the person/assignee/attorney/firm/organization who/which first sends/sent this case to them and by whom/which I hereby declare that I have consented after full disclosure to be represented unless/until I instruct Crawford PLLC.

Please direct all correspondence in this case to Crawford PLLC at the address indicated below:

Crawford PLLC
1270 Northland Drive, Suite 390
St. Paul, Minnesota 55120

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

2	**Full Name Of Inventor**	**Family Name** HANNUKSELA	**First Given Name** MISKA	**Second Given Name**
0	**Residence & Citizenship**	**City** TAMPERE	**State or Foreign Country** FINLAND	**Country of Citizenship** FINLAND
1	**Post Office Address**	**Post Office Address** KUKKANIITYNKATU 4 B	**City** TAMPERE	**State & Zip Code/Country** 33710/FINLAND
Signature of Inventor 201:			**Date:**	7 January 2003

§ 1.56 Duty to disclose information material to patentability.

(a) A patent by its very nature is affected with a public interest. The public interest is best served, and the most effective patent examination occurs when, at the time an application is being examined, the Office is aware of and evaluates the teachings of all information material to patentability. Each individual associated with the filing and prosecution of a patent application has a duty of candor and good faith in dealing with the Office, which includes a duty to disclose to the Office all information known to that individual to be material to patentability as defined in this section. The duty to disclose information exists with respect to each pending claim until the claim is canceled or withdrawn from consideration, or the application becomes abandoned. Information material to the patentability of a claim that is canceled or withdrawn from consideration need not be submitted if the information is not material to the patentability of any claim remaining under consideration in the application. There is no duty to submit information which is not material to the patentability of any existing claim. The duty to disclose all information known to be material to patentability is deemed to be satisfied if all information known to be material to patentability of any claim issued in a patent was cited by the Office or submitted to the Office in the manner prescribed by §§ 1.97(b)-(d) and 1.98. However, no patent will be granted on an application in connection with which fraud on the Office was practiced or attempted or the duty of disclosure was violated through bad faith or intentional misconduct. The Office encourages applicants to carefully examine:

(1) prior art cited in search reports of a foreign patent office in a counterpart application, and

(2) the closest information over which individuals associated with the filing or prosecution of a patent application believe any pending claim patentably defines, to make sure that any material information contained therein is disclosed to the Office.

(b) Under this section, information is material to patentability when it is not cumulative to information already of record or being made of record in the application, and

(1) It establishes, by itself or in combination with other information, a prima facie case of unpatentability of a claim; or

(2) It refutes, or is inconsistent with, a position the applicant takes in:

(i) Opposing an argument of unpatentability relied on by the Office, or

(ii) Asserting an argument of patentability.

- A prima facie case of unpatentability is established when the information compels a conclusion that a claim is unpatentable under the preponderance of evidence, burden-of-proof standard, giving each term in the claim its broadest reasonable construction consistent with the specification, and before any consideration is given to evidence which may be submitted in an attempt to establish a contrary conclusion of patentability.

(c) Individuals associated with the filing or prosecution of a patent application within the meaning of this section are:

(1) Each inventor named in the application:

(2) Each attorney or agent who prepares or prosecutes the application; and

(3) Every other person who is substantively involved in the preparation or prosecution of the application and who is associated with the inventor, with the assignee or with anyone to whom there is an obligation to assign the application.

(d) Individuals other than the attorney, agent or inventor may comply with this section by disclosing information to the attorney, agent, or inventor.

PATENT APPLICATION SERIAL NO. _____

U.S. DEPARTMENT OF COMMERCE
PATENT AND TRADEMARK OFFICE
FEE RECORD SHEET

01/27/2006 HMARZI1 00000049 503581 11338934

01 FC:1011 300.00 DA
02 FC:1111 500.00 DA
03 FC:1311 200.00 DA
04 FC:1201 800.00 DA

PTO-1556
(5/87)

*U.S. Government Printing Office: 2002 — 489-267/69033

PTO/SB/06 (12-04)
Approved for use through 7/31/2006. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

PATENT APPLICATION FEE DETERMINATION RECORD

Substitute for Form PTO-875 Effective December 8, 2004

Application or Docket Number 1133 8931

APPLICATION AS FILED – PART I

FOR	(Column 1) NUMBER FILED	(Column 2) NUMBER EXTRA	SMALL ENTITY RATE ($)	SMALL ENTITY FEE ($)	OR	OTHER THAN SMALL ENTITY RATE ($)	OTHER THAN SMALL ENTITY FEE ($)
BASIC FEE (37 CFR 1.16(a), (b), or (c))	N/A	N/A	N/A	150.00		N/A	300.00
SEARCH FEE (37 CFR 1.16(k), (i), or (m))	N/A	N/A	N/A	$250		N/A	$500
EXAMINATION FEE (37 CFR 1.16(o), (p), or (q))	N/A	N/A	N/A	$100		N/A	$200
TOTAL CLAIMS (37 CFR 1.16(i))	16 minus 20 =	*	X$ 25 =		OR	X$50 =	
INDEPENDENT CLAIMS (37 CFR 1.16(h))	7 minus 3 =	* 4	X100 =			X200 =	
APPLICATION SIZE FEE (37 CFR 1.16(s))	If the specification and drawings exceed 100 sheets of paper, the application size fee due is $250 ($125 for small entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s).						
MULTIPLE DEPENDENT CLAIM PRESENT (37 CFR 1.16(j))			+180=			+360=	
			TOTAL			TOTAL	

* If the difference in column 1 is less than zero, enter "0" in column 2.

APPLICATION AS AMENDED – PART II

AMENDMENT A	(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA	SMALL ENTITY RATE ($)	SMALL ENTITY ADDITIONAL FEE ($)	OR	OTHER THAN SMALL ENTITY RATE ($)	OTHER THAN SMALL ENTITY ADDITIONAL FEE ($)
Total (37 CFR 1.16(i))	*	Minus	**	=	X$ 25 =		OR	X$50 =	
Independent (37 CFR 1.16(h))	*	Minus	***	=	X100 =		OR	X200 =	
Application Size Fee (37 CFR 1.16(s))									
FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))					+180=		OR	+360=	
					TOTAL ADD'L FEE		OR	TOTAL ADD'L FEE	

AMENDMENT B	(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA	RATE ($)	ADDITIONAL FEE ($)		RATE ($)	ADDITIONAL FEE ($)
Total (37 CFR 1.16(i))	*	Minus	**	=	X$ 25 =		OR	X$50 =	
Independent (37 CFR 1.16(h))	*	Minus	***	=	X100 =		OR	X200 =	
Application Size Fee (37 CFR 1.16(s))									
FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))					+180=		OR	+360=	
					TOTAL ADD'L FEE		OR	TOTAL ADD'L FEE	

* If the entry in column 1 is less than the entry in column 2, write "0" in column 3.
** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 20, enter "20".
*** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 3, enter "3".
The "Highest Number Previously Paid For" (Total or Independent) is the highest number found in the appropriate box in column 1.

This collection of information is required by 37 CFR 1.16. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

OIPE IAP2 JUN 19 2006 PATENT & TRADEMARK OFFICE

IFW

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant:	HANNUKSELA	Examiner:	Unassigned
Serial No.:	11/338,934	Group Art Unit:	2613
Filed:	January 25, 2006	Docket No.:	TPO.004.C2
Title:	GROUPING OF IMAGE FRAMES IN VIDEO CODING		

CERTIFICATE UNDER 37 CFR 1.8: The undersigned hereby certifies that this correspondence and the papers, as described hereinabove, are being deposited in the United States Postal Service, as first class mail, in an envelope addressed to: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450 on June 13, 2006.

By Rennae Johnson
Rennae Johnson

INFORMATION DISCLOSURE STATEMENT (37 C.F.R. §1.97(b))

Mail Stop Amendment
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

With regard to the above-identified application, the items of information listed on the enclosed Form 1449 are brought to the attention of the Examiner.

This statement should be considered because it is submitted either within three months of the filing date or before the first Office Action of the above-identified application. Accordingly, no fee is due for consideration of the items listed on the enclosed Form 1449.

In accordance with 37 C.F.R. §1.98(a)(2), and the 5 August 2003 Official Gazette Notice, only a copy of each foreign document or other non-patent published document listed on the enclosed Form 1449 is provided.

Please note that any notations or markings on the attached documents do not reflect particular relevance, or lack thereof, to the present application, nor were they necessarily made by anyone affiliated with the prosecution of the present application.

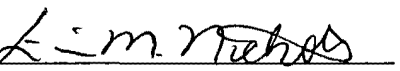
No representation is made that a reference is "prior art" within the meaning of 35 U.S.C. §§102 and 103 and Applicant reserves the right, pursuant to 37 C.F.R. §1.131 or otherwise, to establish that the reference(s) are not "prior art." Moreover, Applicant does not represent that a reference has been thoroughly reviewed or that any relevance of any portion of a reference is intended.

Consideration of the items listed is respectfully requested. Pursuant to the provisions of MPEP §609, it is requested that the Examiner return a copy of the attached Form 1449, marked as being considered and initialed by the Examiner, to the undersigned with the next official communication.

Respectfully submitted,

HOLLINGSWORTH & FUNK, LLC
8009 34th Avenue South, Suite 125
Minneapolis, MN 55425
952.854.2700

Date: June 13, 2006

By: 
Erin M. Nichols
Reg. No. 57,125

OIPE JAN 19 2006 PATENT & TRADEMARK OFFICE

FORM 1449* INFORMATION DISCLOSURE STATEMENT IN AN APPLICATION (Use several sheets if necessary)	Docket Number: TPO.004.C2	Application Number: 11/338,934
	Applicant: HANNUKSELA	
	Filing Date: 01/25/2006	Group Art Unit: 2613

U.S. PATENT DOCUMENTS

EXAMINER INITIAL	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	5,699,476	16.12.1997	VAN DER MEER			
	5,774,593	30.6.1998	ZICK ET AL.			
	6,108,382	22.8.2000	GRINGERI ET AL.			
	6,307,886	23.10.2001	WESTERMANN			
	6,337,881	8.1.2002	CHADDHA			

FOREIGN PATENT DOCUMENTS

	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	TRANSLATION	
						YES	NO
	WO 00/67469	9.11.2000	PCT			n/a	
	WO 01/47283	28.06.2001	PCT			n/a	
	WO 02/054776	11.7.2002	PCT			n/a	
	EP 1 005 232	31.05.2000	EP			n/a	

OTHER DOCUMENTS (Including Author, Title, Date, Pertinent Pages, Etc.)

	9-17 Oct. 2002	Joint Video Team of ISO/IEC MPEG and ITU-T VCEG, *Editor's Proposed Draft Text Modification for Joint Video Specification (ITU-T Rec. 264 ISO/IEC 14496-10 AVC)*, Geneva modifications draft 37.
	15 Jan. 1998	ITU-Telecommunications Standardization Union, *Draft Text of Recommendation H.263 Version 2 ("H.263+") for Decision.*
	7/10/02	ITU-Telecommunications Standardization Union, *H.26L Test Model Long Term Number 8 (TML-8) Draft0*, Document VCEG-N10. pp. 1-46.
	Nov. 2000	ITU-Telecommunications Standardization Union, *Draft for "H.263++" Annexes U,V, and W to Recommendation H.263.*
	12 Nov. 2002	David Singer and Toby Walker, *Study Text of ISO/IEC 14496-15/FCD,* pp. 1-25.
	12 Nov. 2002	David Singer and Toby Walker, *Study Text of ISO/IEC 14496-15/FCD (Revised), 26* pages.
	Sept. 2002	Hannuksela, M et al.; Sub-Picture: ROI Coding and Unequal Error Protection, IEEE 2002 Conference, September 2002, USA (internet publication).
	Sept. 2002	Hannuksela, M et al.; Sub-Picture: Video-Coding for Unequal Error Protection, XI European Signal Processing Conference, September 2002, Toulouse, France (internet publication).
	7.12.2001	Wang, Y.K. et al.; Core Experiment Description of Sub-Picture Coding, ITU 15. Meeting 3-7.12.2001, Pattaya, Thailand (internet publication)
	2001	B.A. Lokshin, "Digital Broadcasting: from Studio to Audience" Cyrus Systems, Moscow, 2001. (translation included)

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609; draw line through citation if not in conformance and not considered. Include copy of this form for next communication to the Applicant.

*Substitute Disclosure Statement Form (PTO-1449) Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) **International Patent Classification:** **H04N 5/00, H04N 7/26**	**A1**	(11) **International Publication Number:**	**WO 00/67469**
		(43) **International Publication Date:**	09 November 2000 (09.11.2000)

(21) **International Application Number:** PCT/FI00/00363

(22) **International Filing Date:** 26 April 2000 (26.04.2000)

(30) **Priority Data:**
990970 29 April 1999 (29.04.1999) FI

(60) **Parent Application or Grant**
NOKIA MOBILE PHONES LTD. [/]; (). HANNUKSELA, Miska [/]; (). WALKER, Andrew ; ().

Published

(54) **Title: DATA TRANSMISSION**
(54) **Titre: TRANSMISSION DE DONNEES**

(57) Abstract

A data transmission system (10) is provided for transmitting a data signal between a server (12, 14, 16, 18) and a communications terminal (20, 22) over a transmission link having a variable bandwidth. The data signal is a scalable compressed multimedia clip comprising a sequence of images. Each image has a base layer and a number of enhancement layers. To cope with variability in the bandwidth, the sequence is re-ordered so that the base layers have greater safety times than the enhancement layers. This allows all of the base layers to arrive in time to be played at the expense of the enhancement layers.

(57) Abrégé

L'invention concerne un système de transmission de données (10) qui assure la transmission d'un signal de données entre un serveur (12, 14, 16, 18) et un terminal de communication (20, 22) sur une liaison à largeur de bande variable. Le signal de données est un signal écrêté multimédia comprimé variable comprenant une séquence d'images. Chaque image comporte une couche de base et un certain nombre de couches d'accentuation. Pour tenir compte de la variabilité de la largeur de bande, on réordonne la séquence pour que la couche de base dispose de temps de sécurité accrus par rapport aux couches d'accentuation. Cela permet aux couches de base de se présenter à temps pour être sollicitées aux dépens des couches d'accentuation.

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : H04N 5/00, 7/26

A1

(11) International Publication Number: **WO 00/67469**

(43) International Publication Date: 9 November 2000 (09.11.00)

(21) International Application Number: PCT/FI00/00363

(22) International Filing Date: 26 April 2000 (26.04.00)

(30) Priority Data:
990970 29 April 1999 (29.04.99) FI

(71) Applicant: NOKIA MOBILE PHONES LTD. [FI/FI]; Keilalahdentie 4, FIN–02150 Espoo (FI).

(72) Inventor: HANNUKSELA, Miska; Kukkaniitynkatu 4 B, FIN–33710 Tampere (FI).

(74) Agent: WALKER, Andrew; Nokia Corporation, P.O. Box 319, FIN–00045 Nokia Group (FI).

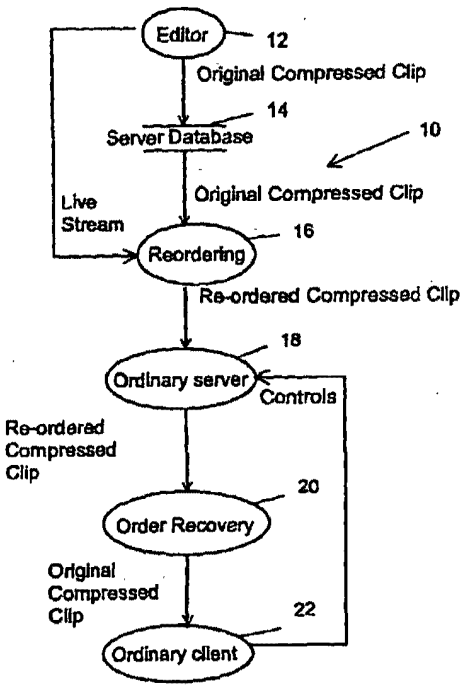
(81) Designated States: AE, AG, AL, AM, AT, AT (Utility model), AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, CZ (Utility model), DE, DE (Utility model), DK, DK (Utility model), DM, DZ, EE, EE (Utility model), ES, FI, FI (Utility model), GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SK (Utility model), SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: DATA TRANSMISSION

(57) Abstract

A data transmission system (10) is provided for transmitting a data signal between a server (12, 14, 16, 18) and a communications terminal (20, 22) over a transmission link having a variable bandwidth. The data signal is a scalable compressed multimedia clip comprising a sequence of images. Each image has a base layer and a number of enhancement layers. To cope with variability in the bandwidth, the sequence is re–ordered so that the base layers have greater safety times than the enhancement layers. This allows all of the base layers to arrive in time to be played at the expense of the enhancement layers.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

Description

5

10

15

20

25

30

35

40

45

50

55

DATA TRANSMISSION

This invention relates to data transmission and is particularly, but not exclusively, related to transmission of data representative of picture sequences, such as video. It is particularly suited to transmission over links susceptible to errors and loss of data.

During the past few years, the amount of multimedia content available through the Internet has increased considerably. Since data delivery rates to mobile terminals are becoming high enough to enable such terminals to be able to retrieve multimedia content, it is becoming desirable to enable mobile terminals to retrieve video and other multimedia content from the Internet. An example of a high-speed data delivery system is the upcoming GSM phase 2+.

The term multimedia as used herein applies to both sound and pictures, to sound only and to pictures only. Sound may include speech and music.

Network traffic through the Internet is based on a transport protocol called the Internet Protocol (IP). IP is concerned with transporting data packets from one location to another. It facilitates the routing of packets through intermediate gateways, that is, it allows data to be sent to machines that are not directly connected in the same physical network. The unit of data transported by the IP layer is called an IP datagram. The delivery service offered by IP is connectionless, that is IP datagrams are routed around the Internet independently of each other. Since no resources are permanently committed within the gateways to any particular connection, the gateways may occasionally have to discard datagrams because of lack of buffer space or other resources. Thus, the delivery service offered by IP is a best effort service rather than a guaranteed service.

Internet multimedia is typically streamed over the User Datagram Protocol (UDP), the Transmission Control Protocol (TCP) or the Hypertext Transfer Protocol (HTTP).

UDP is a connectionless lightweight transport protocol. It offers very little above the service offered by IP. Its most important function is to deliver datagrams between specific transport endpoints. Consequently, the transmitting application has to take care of how to packetize data to datagrams. Headers used in UDP contain a checksum that allows the UDP layer at the receiving end to check the validity of the data. Otherwise, degradation of IP datagrams will in turn affect UDP datagrams. UDP does not check that the datagrams have been received, does not retransmit missing datagrams, nor does it guarantee that the datagrams are received in the same order as they were transmitted.

UDP introduces a relatively stable throughput having a small delay since there are no retransmissions. Therefore it is used in retrieval applications to deal with the effect of network congestion and to reduce delay (and jitter) at the receiving end. However, the client must be able to recover from packet losses and possibly conceal lost content. Even with reconstruction and concealment, the quality of a reconstructed clip suffers somewhat. On the other hand, playback of the clip is likely to happen in real-time without annoying pauses. Firewalls, whether in a company or elsewhere, may forbid the usage of UDP because it is connectionless.

TCP is a connection-orientated transport protocol and the application using it can transmit or receive a series of bytes with no apparent boundaries as in UDP. The TCP layer divides the byte stream into packets, sends the packets over an IP network and ensures that the packets are error-free and received in their correct order. The basic idea of how TCP works is as follows. Each time TCP sends a packet of data, it starts a timer. When the receiving end gets the packet, it immediately sends an acknowledgement back to the sender. When the sender receives the acknowledgement, it knows all is well, and cancels the timer. However, if the IP layer loses the outgoing segment or the return acknowledgement, the timer at the sending end will expire. At this point, the sender will retransmit the segment. Now, if the sender waited for an acknowledgement for each packet before sending the next one, the overall

transmission time would be relatively long and dependent on the round-trip delay between the sender and the receiver. To overcome this problem, TCP uses a sliding window protocol that allows several unacknowledged packets to be present in the network. In this protocol, an acknowledgement packet contains a field filled with the number of bytes the client is willing to accept (beyond the ones that are currently acknowledged). This window size field indicates the amount of buffer space available at the client for storage of incoming data. The sender may transmit data within the limit indicated by the latest received window size field. The sliding window protocol means that TCP effectively has a slow start mechanism. At the beginning of a connection, the very first packet has to be acknowledged before the sender can send the next one. Typically, the client then increases the window size exponentially. However, if there is congestion in the network, the window size is decreased (in order to avoid congestion and to avoid receive buffer overflow). The details how the window size is changed depend on the particular TCP implementation in use.

A multimedia content creation and retrieval system is shown in Figure 1. The system has one or more media sources, for example a camera and a microphone. Alternatively, multimedia content can also be synthetically created without a natural media source, for example animated computer graphics and digitally generated music. In order to compose a multimedia clip consisting of different media types, such as video, audio, text, images, graphics and animation, raw data captured from the sources are edited by an editor. Typically the storage space taken up by raw (uncompressed) multimedia data is huge. It can be megabytes for a video sequence which can include a mixture of different media, for example animation. In order to provide an attractive multimedia retrieval service over low bit rate channels, for example 28.8 kbps and 56 kbps, multimedia clips are compressed in the editing phase. This typically occurs off-line. The clips are then handed to a multimedia server. Typically, a number of clients can access the server over one or more networks. The server is able to respond to the requests presented by the clients. The main task of the server is to transmit a desired multimedia clip to the client which the client decompresses and plays. During

playback, the client utilizes one or more output devices, such as a screen and a loudspeaker. In some circumstances, clients are able to start playback while data are still being downloaded.

It is convenient to deliver a clip by using a single channel which provides a similar quality of service for the entire clip. Alternatively different channels can be used to deliver different parts of a clip, for example sound on one channel and pictures on another. Different channels may provide different qualities of service. In this context, quality of service includes bit rate, loss or bit error rate and transmission delay variation.

In order to ensure multimedia content of a sufficient quality is delivered, it is provided over a reliable network connection, such as TCP, which ensures that received data are error-free and in the correct order. Lost or corrupted protocol data units are retransmitted. Consequently, the channel throughput can vary significantly. This can even cause pauses in the playback of a multimedia stream whilst lost or corrupted data are retransmitted. Pauses in multimedia playback are annoying.

Sometimes retransmission of lost data is not handled by the transport protocol but rather by some higher-level protocol. Such a protocol can select the most vital lost parts of a multimedia stream and request the retransmission of those. The most vital parts can be used for prediction of other parts of the stream, for example.

In order to understand the invention better, descriptions of the elements of the retrieval system, namely the editor, the server and the client, are set out below.

A typical sequence of operations carried out by the multimedia clip editor is shown in Figure 2. Raw data are captured from one or more data sources. Capturing is done using hardware, device drivers dedicated to the hardware and a capturing application which controls the device drivers to use the hardware. Capturing hardware may consist of a video camera connected to a PC video grabber card,

for example. The output of the capturing phase is usually either uncompressed data or slightly compressed data with irrelevant quality degradations when compared to uncompressed data. For example, the output of a video grabber card could be in an uncompressed YUV 4:2:0 format or in a motion-JPEG format. The YUV colour model and the possible sub-sampling schemes are defined in Recommendation ITU-R BT.601-5 "Studio Encoding Parameters of Digital Television for Standard 4:3 and Wide-Screen 16:9 Aspect Ratios". Relevant digital picture formats such as CIF, QCIF and SQCIF are defined in Recommendation ITU-T H.261 "Video Codec for Audiovisual Services at p x 64 kbits" (section 3.1 "Source Formats").

During editing separate media tracks are tied together in a single timeline. It is also possible to edit the media tracks in various ways, for example to reduce the video frame rate. Each media track may be compressed. For example, the uncompressed YUV 4:2:0 video track could be compressed using ITU-T recommendation H.263 for low bit rate video coding. If the compressed media tracks are multiplexed, they are interleaved so that they form a single bitstream. This clip is then handed to the multimedia server. Multiplexing is not essential to provide a bitstream. For example, different media components such as sounds and images may be identified with packet header information in the transport layer. Different UDP port numbers can be used for different media components.

A typical sequence of operations carried out by the multimedia server is shown in Figure 3. Typically multimedia servers have two modes of operation; they deliver either pre-stored multimedia clips or a live (real-time) multimedia stream. In the first mode, clips are stored in a server database, which is then accessed on-demand by the server. In the second mode, multimedia clips are handed to the server as a continuous media stream that is immediately transmitted to clients. Clients control the operation of the server by an appropriate control protocol being at least able to select a desired media clip. In addition, servers may support more advanced controls. For example, clients may be able to stop the transmission of a clip, to pause and resume transmission of a clip, and to control the media flow in

case of a varying throughput of the transmission channel in which case the server must dynamically adjust the bitstream to fit into the available bandwidth.

A typical sequence of operations carried out by the multimedia retrieval client is shown in Figure 4. The client gets a compressed and multiplexed media clip from a multimedia server. The client demultiplexes the clip in order to obtain separate media tracks. These media tracks are then decompressed to provide reconstructed media tracks which are played out with output devices. In addition to these operations, a controller unit is provided to interface with end-users, that is to control playback according to end-user input and to handle client-server control traffic. It should be noted that the demultiplexing-decompression-playback chain can be done on a first part of the clip while still downloading a subsequent part of the clip. This is commonly referred to as streaming. An alternative to streaming is to download the whole clip to the client and then demultiplex it, decompress it and play it.

A typical approach to the problem of varying throughput of a channel is to buffer media data in the client before starting the playback and/or to adjust the transmitted bit rate in real-time according to channel throughput statistics.

One way of tackling the problem of pauses is by using dynamic bit rate adjustment in the multimedia server. However, the server may not react to network congestion sufficiently quickly to avoid pausing in the client. In addition, the server cannot control the retransmission mechanism of TCP (or other underlying protocols, such as IP).

Even if dynamic bit rate adjustment is used, the client has to do some initial buffering in any case to avoid delivery delays caused by re-transmission. If a constant channel bit rate is assumed, one can calculate the point in time at which a data unit is supposed to have been completely received. In addition, one can calculate the point in time by which a data unit is supposed to have been played. The time difference between these two points in time is referred to as the safety

time. Another way of defining the safety time is to state that it is the maximum time between two consecutively received data units which does not cause pausing in the playback.

When calculating the safety times for a clip, each data unit has to be considered separately. The calculations assume that no throughput drops occur before the data unit that is currently being processed. If the maximum throughput of the channel is equal to the average bit rate of the multimedia clip, the client cannot recover from a drop in the amount of received bits after throughput has dropped. The only way to guarantee some protection against throughput drops is to buffer some data before starting playback. If the channel stops providing data, the client can still play the stream while there are data in a buffer. Thus, the average safety time is approximately equal to the initial buffering time. Since the bit rate of the clip may vary, the safety time also may vary and the minimum safety time is thus equal to or less than the initial buffering delay.

HTTP, the Hypertext Transfer Protocol, is the basis of the World Wide Web (WWW). It is a simple protocol. A client establishes a TCP connection to a server, issues a request, and reads back the server's response. The server denotes the end of its response by closing the connection. The arrangement of protocol layers is typically HTTP on TCP which is on IP.

The most common HTTP request is called GET. The GET request is associated with a universal resource identifier (URI) which uniquely specifies the requested item. The server responds to the GET request by returning the file corresponding to the specified URI. The file returned by the server normally contains pointers (hypertext links) to other files that can reside on other servers. A user can therefore easily follow the links from file to file.

Servers used for Internet multimedia retrieval are either dedicated multimedia servers or normal WWW servers.

Dedicated multimedia servers are typically capable of transmission over HTTP, TCP and UDP protocols. They may be able to readjust the contents of media clips dynamically to meet the available network bandwidth and to avoid network congestion. They may also support fast forward and fast rewind operations as well as live multimedia streaming. They can provide a number of streams simultaneously.

Multimedia servers based on normal WWW servers are also referred to as server-less or HTTP multimedia solutions. The multimedia clips are streamed over HTTP. Since this type of server has no control over the contents of the stream, no flow (bandwidth) control can be applied, and it cannot respond to network congestion. Therefore, sudden pauses in the playback can occur. Consequently there must be a relatively long initial buffering delay in the client before starting the playback to avoid such sudden pauses. Fast-forwarding a multimedia stream from a standard WWW server is not possible. Live multimedia streaming must be implemented using special tricks, such as Java programming.

When a streamed multimedia clip is received, a suitable independent media player application or a browser plug-in can be used to play it. Such multimedia players differ largely from browser to browser. Newer browsers may have some integrated plug-ins for the most popular streaming video players.

There are a number of different data transmission methods available for transmitting data between mobile terminals and their networks. One of the best known methods is GSM (Global System for Mobile communications).

The current GSM data service called Circuit Switched Data (CSD) offers a 9.6 kbps circuit-switched channel. It is intended for GSM to provide a 14.4 kbps data channel having forward error correction (FEC) and status information. High Speed Circuit Switched Data (HSCSD) provides multiple 9.6 kbps or 14.4 kbps time slots for a single user at the same time. There are symmetric and asymmetric connections. In a symmetric connection, the air interface resources are allocated

symmetrically, and provide the same data transmission rate in both directions. In an asymmetric connection, different data rates are supported for up-link and for down-link. However, asymmetric air interface connections are applicable only in non-transparent mode (see below).

Circuit-switched GSM data systems, CSD and HSCSD, offer two basic connection types, namely transparent (T) and non-transparent (NT). These are distinguished by the way they correct transmission errors. In a transparent connection, error correction is done solely by a forward error correction mechanism provided by the radio interface transmission scheme. The connection is considered as a synchronous circuit. The available throughput is constant, and the transmission delay is fixed. The transmitted data are likely to contain bit inversion errors. In a non-transparent connection, the GSM circuit connection is regarded as a packet (or frame) data flow although the end-to-end service is circuit-switched. Each frame includes redundancy bits to enable a receiver to detect remaining errors. There are two sources of error, packet drops and corruption due to interference in the radio frequency path. The latter can be recovered by redundancy checking. The Radio Link Protocol (RLP) is used to provide retransmission in case of remaining errors in frame. If a frame is found to be correct, the receiver acknowledges this fact. If it is found not to be correct, a negative acknowledgement is sent and the indicated frame is retransmitted. Consequently, a non-transparent connection is error-free but throughput and transmission delay vary.

Other network types exist, for example GPRS (General Packet Radio System). In GPRS the transmissions are truly packet based.

A video sequence consists of a series of still pictures. Video compression methods are based on reducing redundant and perceptually irrelevant parts of video sequences. The redundancy in video sequences can be categorised into spatial, temporal and spectral redundancy. Spatial redundancy means the correlation between neighbouring pixels. Temporal redundancy means that the same objects

appear in consecutive pictures. Reducing the temporal redundancy reduces the amount of data required to represent a particular image sequence and thus compresses the data. This can be achieved by generating motion compensation data, which describe the motion between the current and a previous (reference) picture. In effect, the current picture is predicted from the previous one. Spectral redundancy means the correlation between the different colour components of the same picture.

Simply reducing the redundancy of the sequence does not usually compress it enough. Therefore, some video encoders try to reduce the quality of those parts of a video sequence which are subjectively the least important. In addition, the redundancy of the encoded bitstream is reduced by means of efficient lossless coding of compression parameters and coefficients. The main technique is to use variable length codes.

Video compression methods typically differentiate between pictures that can use temporal redundancy reduction and those that cannot. Compressed pictures, which do not use temporal redundancy reduction methods, are usually called INTRA or I-frames whereas temporally predicted pictures are called INTER or P-frames. In the INTER frame case, the predicted (motion-compensated) picture is rarely precise enough, and therefore a spatially compressed prediction error picture is also associated with each INTER frame.

Temporal scalability provides a mechanism for enhancing perceptual quality by increasing the picture display rate. This is achieved by taking a pair of consecutive reference pictures and bi-directionally predicting a B-picture from either one or both of them. The B-picture can then be played in sequence between the two anchor pictures. This is illustrated in Figure 5. Bi-directional temporal prediction yields a more accurately predicted picture than uni-directional prediction. Thus, for the same quantization level, B-pictures yield increased compression as compared to forwardly predicted P-pictures. B-pictures are not used as reference pictures, that is other pictures are never predicted from them. Since they can be discarded

without impacting the picture quality of future pictures, they provide temporal scalability. It should be noted that while B-pictures provide better compression performance than P-pictures, they are more complex to construct and require more memory. Furthermore they introduce additional delays because bi-directional interpolation requires both reference pictures to have been received and additional calculations are required. In addition, B-pictures require more side information in the bitstream.

The term scalability refers to the capability of a compressed sequence to be decoded at different data rates. In other words, a scalable multimedia clip can be edited relatively easily while it is compressed so that it can be streamed over channels with different bandwidths and still be decoded and played back in real-time.

Scalable multimedia is typically ordered so that there are hierarchical layers of data. A base layer contains a basic representation of the multimedia clip whereas enhancement layers contain refinement data on top of underlying layers. Consequently, the enhancement layers improve the quality of the clip.

Scalability is a desirable property for heterogeneous and error prone environments. This property is desirable in order to counter limitations such as constraints on bit rate, display resolution, network throughput, and decoder complexity.

Scalability can be used to improve error resilience in a transport system where layered coding is combined with transport prioritisation. The term transport prioritisation here refers to various mechanisms to provide different qualities of service in transport, including unequal error protection, to provide different channels having different error/loss rates. Depending on their nature, data are assigned differently, for example, the base layer may be delivered through a channel with high degree of error protection, and the enhancement layers may be transmitted through more error-prone channels.

Generally, scalable multimedia coding suffers from a worse compression efficiency than non-scalable coding. In other words, a multimedia clip coded as a scalable multimedia clip with all enhancement layers requires greater bandwidth than if it had been coded as a non-scalable single-layer clip with equal quality. However, exceptions to this general rule exist, for example the temporally scalable B-frames in video compression.

In the following, scalability is discussed with reference to the ITU-T H.263 video compression standard. H.263 is an ITU-T recommendation for video coding in low bit rate communication which generally means data rates below 64 kbps. The recommendation specifies the bitstream syntax and the decoding of the bitstream. Currently, there are two versions of H.263. Version 1 consists of the core algorithm and four optional coding modes. H.263 version 2 is an extension of version 1 providing twelve new negotiable coding modes.

Pictures are coded as luminance and two colour difference (chrominance) components (Y, C_B and C_R). The chrominance pictures are sampled with a half of the pixels along the both co-ordinate axes when compared to the luminance picture.

The scalability mode (Annex O) of H.263 specifies syntax to support temporal, signal-to noise ratio (SNR), and spatial scalability capabilities.

Spatial scalability and SNR scalability are closely related, the only difference being the increased spatial resolution provided by spatial scalability. An example of SNR scalable pictures is shown in Figure 6. SNR scalability implies the creation of multi-rate bit streams. It allows for the recovery of coding errors, or differences between an original picture and its reconstruction. This is achieved by using a finer quantizer to encode the difference picture in an enhancement layer. This additional information increases the SNR of the overall reproduced picture.

Spatial scalability allows for the creation of multi-resolution bit streams to meet varying display requirements and/or constraints. A spatially scalable structure is illustrated in Figure 7. It is essentially the same as in SNR scalability except that a spatial enhancement layer attempts to recover the coding loss between an up-sampled version of the reconstructed reference layer picture and a higher resolution version of the original picture. For example, if the reference layer has a quarter common intermediate format (QCIF) resolution, and the enhancement layer has a common intermediate format (CIF) resolution, the reference layer picture must be scaled accordingly such that the enhancement layer picture can be predicted from it. The QCIF standard allows the resolution to be increased by a factor of two in the vertical direction only, horizontal direction only, or both the vertical and horizontal directions for a single enhancement layer. There can be multiple enhancement layers, each increasing the picture resolution over that of the previous layer. The interpolation filters used to up-sample the reference layer picture are explicitly defined in the H.263 standard. Aside from the up-sampling process from the reference to the enhancement layer, the processing and syntax of a spatially scaled picture are identical to those of an SNR scaled picture.

In either SNR or spatial scalability, the enhancement layer pictures are referred to as EI- or EP-pictures. If the enhancement layer picture is upwardly predicted from a picture in the reference layer, then the enhancement layer picture is referred to as an Enhancement-I (EI) picture. In this type of scalability, the reference layer means the layer "below" the current enhancement layer. In some cases, when reference layer pictures are poorly predicted, over-coding of static parts of the picture can occur in the enhancement layer, causing an unnecessarily excessive bit rate. To avoid this problem, forward prediction is permitted in the enhancement layer. A picture that can be predicted in the forward direction from a previous enhancement layer picture or, alternatively, upwardly predicted from the reference layer picture is referred to as an Enhancement-P (EP) picture. Note that computing the average of the upwardly and forwardly predicted pictures can provide bi-directional prediction for EP-pictures. For both EI- and EP-pictures, upward prediction from the reference layer picture implies that no motion vectors are

required. In the case of forward prediction for EP-pictures, motion vectors are required.

In multi-point and broadcast multimedia applications, constraints on network throughput may not be foreseen at the time of encoding. Thus, a scalable bitstream should be used. Figure 8 shows an IP multicasting arrangement where each router can strip the bitstream according to its capabilities. It shows a server S providing a bitstream to a number of clients C. The bitstreams are routed to the clients by routers R. In this example, the server is providing a clip which can be scaled to at least three bit rates, 120 kbit/s, 60 kbit/s and 28 kbit/s.

If the client and server are connected via a normal uni-cast connection, the server may try to adjust the bit rate of the transmitted multimedia clip according to the temporary channel throughput. One solution is to use a layered bit stream and to adapt to bandwidth changes by varying the number of transmitted enhancement layers.

According to a first aspect of the invention there is provided a method of transmitting a data signal having a sequence of data units in a predetermined order over a transmission link between a data source and a data sink comprising the step of sending the data units in an order determined by their relative importance rather than their predetermined order.

According to a second aspect of the invention there is provided a data transmission system for transmitting a data signal having a sequence of data units in a predetermined order over a transmission link between a data source and a data sink the data units being sent in an order determined by their relative importance rather than their predetermined order.

According to a third aspect of the invention there is provided a server for transmitting a data signal having a sequence of data units in a predetermined

order over a transmission link the data units being sent in an order determined by their relative importance rather than their predetermined order.

Preferably the server comprises re-ordering means. Preferably the server comprises an editor for providing a scalable data signal.

According to a fourth aspect of the invention there is provided a computer program product stored on a computer usable medium comprising computer readable program means for causing transmission of a data signal having a sequence of data units in a predetermined order over a transmission link the data units being sent in an order determined by their relative importance rather than their predetermined order.

Preferably the computer program product comprises a server. In this sense it is meant that the when the computer program is loaded into hardware it causes the hardware to operate as a server device. Preferably the computer program product comprises an editor for providing a scalable data signal. Preferably the computer program product comprises re-ordering means for re-ordering the data units.

According to a fifth aspect of the invention there is provided a data signal having a sequence of data units for transmission over a transmission link between a data source and a data sink the data units being in an order determined by their relative importance rather than their predetermined order.

Preferably where the data units or layers of the data units are ordered into a sequence other than their original sequence, the invention also includes means or a step for returning the data units or layers of the data units into their original sequence.

The transmission link may simply be a bottleneck where congestion is occurring and the invention is used to overcome the bottleneck.

An advantage of the invention is that it can provide larger safety times than prior art arrangements with the same initial buffering. In addition, it provides some degree of protection against decreased transmission throughput. Therefore it helps in reducing the occurrence of pauses without sacrificing compression efficiency.

The invention also helps in transmitting over a unreliable network connection using transmission protocols which do not automatically retransmit but only retransmit those missing parts chosen by a higher level function, for example in response to feedback. In this case, it provides more time for retransmission of the vital parts of a data signal than prior art arrangements. Thus, vital parts of a stream are more likely to be transmitted successfully and the ultimate end user of the data signal is likely to perceive a better output quality.

Preferably the bandwidth of the transmission link available to the data signal is variable. This may be because the amount of data traffic using the link is variable, for example due to the need to re-transmit lost or corrupted data.

The invention provides a way of altering a data signal so that it can be transmitted over a transmission link having a bandwidth which would otherwise not be sufficient for it or for allowing certain of the bandwidth of its transmission link to be made available for other data, for example control information.

Preferably the effect of the invention is to increase the safety times of some data units relative to, or even at the expense of, the safety times of other data units.

Preferably the data signal is arranged in a hierarchy having a base layer and at least one enhancement layer. The term base layer means a most fundamental layer. The base layer may represent the least refined version of the data signal. In each unit, the (or each) enhancement layer, when combined with the base layer, refines the base layer to make it more like the data unit. Each enhancement layer

may add more refinement to the base layer. Preferably the safety time of the base layer is increased relative to the safety times of the enhancement layers.

Preferably the data signal is scalable. The base layer and the or each enhancement layer may have a scaled relationship with respect to each other. The signal may be scalable in the temporal, the spatial, the spectral or the SNR domains.

Preferably by virtue of re-ordering the sequence the layers have safety times which depend on their position in the hierarchy. For example the base layer of a particular data unit may have a greater safety time than enhancement layers. Advantageously, the finer the level of refinement represented by a layer, the shorter its safety time.

Preferably the source is a server. Alternatively it is an editor.

Preferably the sink is a client. It may be a destination of the data signal such as a mobile terminal, for example a mobile telephone.

Preferably the data units are frames. In one embodiment the data units represent a sequence of pictures. Conveniently it is a video sequence. Most preferably, it comprises multimedia data.

It should be understood that the terms "layer" and "layers" are being used figuratively and refer to blocks of data which can be applied to each other rather than to a strictly positional relationship.

The invention will now be described by way of example only with reference to the accompanying drawings in which:

Figure 1 shows a multimedia content creation and retrieval system;

Figure 2 shows operations of a multimedia clip editor;

Figure 3 shows operations of a multimedia server;

Figure 4 shows operations of a multimedia retrieval client;

Figure 5 shows an illustration of B picture prediction;

Figure 6 shows an illustration of SNR scalability;

Figure 7 shows an illustration of spatial scalability;

Figure 8 shows scalable multimedia being used in an IP multicast system;

Figure 9 shows mobile multimedia retrieval system according to the invention;

Figure 10 shows operations of an editor according to the invention;

Figure 11 shows operations of a client according to the invention;

Figure 12 shows safety times of a re-ordered sequence and an original sequence;

Figure 13 shows receiving and playback of a re-ordered sequence;

Figure 14 also shows receiving and playback of a re-ordered sequence; and

Figure 15 shows another multimedia content creation and retrieval system.

Figures 1 to 8 have been discussed above.

A mobile multimedia retrieval system 10 according to the invention is shown in Figure 9. Although only a single client 22 is shown it will be appreciated that there can be, and in preferred embodiments will be, a number of such clients. It will also be appreciated that at least part of the transmission link connecting a server 18 and the client 22 will be wireless.

The system 10 comprises an editor 12 which produces a compressed clip, a server database 14 for storing the compressed clip, re-ordering means 16 for re-ordering the sequence of the compressed clip, a server 18 for sending the re-ordered compressed clip and a client comprising order recovery means 20 to recover the original sequence of the compressed clip. The client 22 controls the operation of the server, for example by specifying which data are sent and how they are sent. Clearly, the re-ordering means 16 and the order recovery means 20 need to be compatible so that the latter can recover the original order. The re-ordering means 16 could be located to operate on data before they reach the server database 14 rather than the embodiment shown in Figure 9. Certain types of data, such as a live stream, may by-pass the database 14 and be sent by the

editor 12 directly to the re-ordering means 16. Although they are shown separately in Figure 9, the editor 12, the server database 14, the re-ordering means 16 and the server 18 can all be configured into a single unit or entity which can also be referred to as a server.

Figure 10 shows operation of an editor according to the invention. Raw data captured in the form of one or more uncompressed streams are edited to produce one or more uncompressed tracks. These tracks are compressed in a compression step to produce compressed tracks. The compressed tracks are multiplexed together to produce a compressed multimedia clip. The multimedia clips are coded scalably. The resulting bitstream consists of basic units, such as frames in video compression. These basic units are not all of equal importance; some are primary basic units which are vital to decode the bitstream and some are secondary basic units which improve the quality of the bitstream. Elimination of the secondary basic units causes only a quality loss in the decoded multimedia clip. In effect, the primary basic units are a base layer of the bitstream and the secondary units are enhancement layers.

Instead of organising the basic units into the original transmission order, that is the sequence in which they are produced by editing and compression of raw data, the invented method re-orders their sequence by prioritising the primary basic units so that they pre-empt the secondary basic units in the transmission order. Reordering can take place before or after the multiplexing process depending on the nature of the system and the data. In Figure 10 it is shown as taking place after multiplexing.

A re-ordering algorithm is used to re-order the sequence. In a preferred embodiment, the algorithm conducts stream analysis to identify primary and secondary basic units and to collect bit rate statistics. These may include the number of bits per frame. The bitstream is then re-ordered with iterative re-ordering steps to minimise the initial buffering time needed for full-quality playback with full channel throughput and reduced-quality playback with a reduced channel throughput. An example of such a re-ordering algorithm is given below. It should

be noted that the expected reduced channel throughput has to be estimated in advance since it is one of the input parameters of the algorithm. The estimation can be based on the statistics obtained from the channel behaviour, for example.

Re-ordering of the basic units means that the primary basic units are likely to be received earlier than indicated in their playback time even though some decrease in the channel throughput occurs. However, secondary basic units are vulnerable to transmission delay variations since the method schedules them to be received just before their indicated playback time. In this way the client is still likely to be able to play the primary basic units without pausing the playback, even if the channel throughput decreases considerably. The secondary basic units, on the other hand, may arrive later than their scheduled playback time and will not be used in playback.

Figure 11 shows operation of a client according to the invention. The bitstream is received by the client, re-ordered to produce the original compressed clip, demultiplexed to produce the compressed tracks and then decompressed to reconstruct the tracks. The basic unit order recovery step in the client can be applied either before or after demultiplexing (depending in when it occurs in the content editor). In Figure 11 it is shown as occurring before demultiplexing, to correspond to the order described in Figure 10. In this way re-ordering and order recovery are both applied to the multiplexed bitstream. Order recovery organises the basic units back into their natural sequence for decompression and discards all late-arriving secondary basic units.

On receipt of a basic unit by the client, it is time-stamped with a time by which it should be played. This time-stamp is compared with the current playback time. If the time-stamp is greater than the current playback time the basic unit is not decompressed and it is discarded. If the time-stamp is less than the current playback time the basic unit is decompressed and the two times are compared after decompression. If the time stamp is less than the current playback time the basic unit is played. Otherwise, it is discarded.

In the particular case in which re-ordering is performed after multiplexing, the multiplexing/demultiplexing format used needs to allow re-ordering of the basic units and recovery of the original order.

Figure 15 shows another view of a multimedia content creation and retrieval system according to the invention. The system 50 comprises an editor 52, a server 54 and a client 56.

The editor comprises a compressor 58 for compressing a raw data stream typically being in the form of a sequence of images into a form suitable for transmission, a multiplexer 60 for multiplexing the compressed raw data stream with other compressed data streams together to produce a compressed bitstream and a layer generation means 62 for producing primary basic frames which are vital to decode the bitstream and secondary basic frames which improve the quality of the bitstream.

The server 54 comprises re-ordering means 64 for arranging the frames in an order which depends on their importance in helping to reconstruct the raw data stream at the client 56 rather than their natural sequential order.

The server 54 and the client 56 each comprise transmitting/receiving means 66 and 70 respectively for data transmission over a transmission link 68 between the server 54 and the client 56.

The client 56 comprises order recovery means 72 for putting the frames back into their natural sequential order, a demultiplexer 74 for demultiplexing the bitstream and a decompressor 76 for decompressing the compressed data streams to reproduce the raw data stream. The client comprises playing/output means 78 for playing or otherwise outputting the raw data stream. The client also comprises a controller 80 which applies a control signal to the server 54 to control the re-ordering means in response to receipt of frames in the client 56.

Example

An embodiment of the invention was tested using an H.263 compliant encoder and a frame re-ordering tool. The encoder was capable of producing B-frames. The re-ordering tool comprised a re-ordering algorithm and could be adjusted to produce frame ordering for different decreased channel data rates.

A standard ITU-T test sequence of video frames called Glasgow was used. The sequence had a QCIF resolution, that is a luminance picture size of 176 x 144 pixels and a chrominance picture size of 88 x 72 pixels.

The first 297 frames of the Glasgow sequence were compressed to produce a temporally scalable bitstream comprising INTRA (I) frames, INTER (P) frames and bi-directionally predicted (B) frames. The I and P frames are referred to as primary frames and the B frames are referred to as secondary frames. The frame rate of the original (default) sequence was 12.5 frames per second (fps) and so the compressed sequence corresponded to 23.68 seconds. Each frame was compressed. The first frame was an INTRA frame, and the second one was an INTER frame. After that, there were alternate INTER frames and B-frames. A constant quantizer was used to produce a bitstream having an average bit rate of 41073 bps consisting of an average bit rate for the primary frames of 28984 bps and an average bit rate for the secondary frames of 12089 bps.

For the sake of simplicity, the example does not include multiplexing of the bitstream with another bitstream.

Re-ordering of the frame sequence was optimised separately for bit rates corresponding to sums of the primary frame bit rate (A) and (i) 25%; (ii) 50%; and (iii) 75 % of the secondary frame bit rate (B). It should be noted that 100% of the secondary bit rate (B) corresponds to the average bit rate required in order to obtain full quality playback. The summed bit rates were 32007, 35029 and 38051 bps respectively. For convenience, in the following description, these bit rates are

referred to as 25% (A+0.25B), 50% (A+0.5B) and 75% (A+0.75B), and the corresponding re-ordered sequences are referred to as 25-ordered, 50-ordered and 75-ordered. It was expected that the full rate of the channel would be the average bit rate of the sequence.

The reordering algorithm used two values:
FQFR (full quality full rate), which is the minimum initial buffering time to ensure full-quality playback assuming a full rate channel; and
RQRR (reduced quality reduced rate) which is the minimum initial buffering time to ensure reduced-quality playback in which the primary frames are played and some or all secondary frames may not be played (assuming a reduced secondary bit rate of 25%, 50% or 75%).

In calculating FQFR and RQRR for a particular channel, the algorithm assumes that it has a constant bit rate. Using the bit rate statistics the likely number of bits per frame is calculated. The time instant at which a particular frame is to be played is known from the time stamp. A time is determined by which a frame will be ready for display as a picture (allowing, of course, for the time needed to receive the bits for a frame and to prepare them) and this time is compared with the time stamp of the frame. If the frame will be received after it is to be displayed, the buffering time is increased, if it is received in time then the next frame can be dealt with. Once the algorithm has compared the likely times of receipt of all of the frames against their playback times and adjusted to buffering time accordingly so that they should all be received in time, the minimum buffering time has been calculated.

The algorithm was applied to a sequence having its original order. In this case, FQFR is smaller than RQRR because in order to obtain full quality playback, you have to wait longer in starting playback if you have a reduced rate channel. The algorithm seeks to minimise the initial buffering time needed for full-quality playback with full channel throughput and at the same time for reduced-quality playback with a reduced channel throughput. In other words, the algorithm seeks to find a value where FQFR and RQRR are approximately equal. It should be

noted that the algorithm is not optimal, that is the algorithm does not necessarily re-order the sequence so that the minimum initial buffering delay is found. The algorithm consists of the following steps:

1. Calculate FQFR and RQRR.

2. If FQFR is greater than or equal to RQRR, stop re-ordering. The minimum initial buffering delay is the greater of FQFR or RQRR. Otherwise, continue with the next step.

3. Shift all primary frames towards the beginning of the sequence by one frame step. For example, if the initial sequence is:

I0, P1, P2, B3, P4, B5, P6, B7, P8, B9, P10, ...,

shifting changes the sequence to:

I0, P1, P2, P4, B3, P6, B5, P8, B7, P10, B9, ...

where I0 is the first frame. It can be seen that where a secondary frame is earlier in sequence than an adjacent primary frame, the positions of the primary and secondary frames are exchanged.

4. Continue from step 1.

The buffering time needed to ensure pause-less playback in the reduced bit rate channel is greater for the original sequence than for the re-ordered sequence. It is desirable to have a short initial buffering time. The re-ordering process makes a trade-off between the minimum buffering time needed for a full-rate channel and the minimum buffering time needed for expected reduced-rate channels. Without re-ordering, less initial buffering time would have been needed to provide full-quality playback in the full-rate channel. However, without re-ordering, more initial buffering time would have been required to provide reduced-quality playback in the reduced-rate channel.

In the example, safety times were also investigated. Initial buffering time was set to be the minimum for reduced quality playback of the re-ordered sequence at 25% secondary bit rate. Figure 12 shows how safety time varies with playback of the 25-ordered sequence and the original order sequence. It shows that throughout playback, the 25-ordered sequence (represented by the upper line) has a longer safety time than the original order sequence.

Results were also obtained for the 50% and 75% secondary bit rates. The average safety times (in seconds) for all primary frames in a sequence are shown in the following table for all three reduced bit rates. The safety times are characterised by minimum safety time and average safety time.

		25%	50%	75%
Minimum	Re-ordered	4.79	2.51	1.06
	Default	3.46	1.63	0.64
	Difference	1.33	0.88	0.42
Average	Re-ordered	5.90	3.51	2.12
	Default	4.66	2.83	1.84
	Difference	1.24	0.67	0.28

It can be seen that both the minimum and average safety times of the re-ordered sequences are longer than those of the original sequence.

In contrast, set out below is a table showing the required buffering delays in seconds for full-quality playback at different average bit rates. This is to minimise pauses in playback.

	25%	50%	75%
Re-ordered	10.04	6.03	3.53
Default	7.02	4.64	3.02

It can be seen that the re-ordered sequence needs a longer buffering delay for full-quality playback because the secondary frames are transmitted out of their natural order and time is required to put all the frames back into their natural order. However, this is not important because the invention sets out to provide reduced quality playback with minimal initial buffering delay. To illustrate this, the following

table shows the buffering delays required to ensure pause-less reduced quality playback:

	25%	50%	75%
Re-ordered	5.28	3.49	2.47
Default	6.91	4.46	2.84

These results show that where the bit rate of the transmission channel is reduced, the invented method requires a shorter buffering delay before playback can be started.

The effect of the invention is illustrated in Figure 13 which shows playback of the 25-ordered sequence at 25% bit rate (that is A+0.25B as discussed above). Playback times, that is the times at which the frames are intended or supposed to be played to a viewer, are represented as a straight line. To be played in sequence, without requiring pauses, each frame must be received before its playback time. Any symbol located below the playback time line represents a frame which is received before its scheduled playback time and any symbol located above the playback time line represents a frame which is received after its scheduled playback time. In the example of Figure 13, the primary frames were all received in time to be played and the secondary frames were all received too late. Therefore the secondary frames were all discarded without being played.

It should be understood that if the sequence had not been re-ordered, there is a likelihood that some of the primary frames would have arrived after their playback times and so would have required playback to be paused (since a primary frame cannot be discarded without also discarding all of the frames predicted from it).

Figure 14 shows how the 25-ordered sequence is received and played when the channel throughput is not reduced. It can be seen that all frames are received in time to be played. Again, the upper line represents when secondary frames are received and the lower line represents when primary frames are received.

Of course, the exact appearance of Figures 13 and 14 depends on the nature of the original sequence, as well as the bit rate and the re-ordering that is performed. Differences can be expected in other systems which re-order data in different ways.

In use, the invention can be applied on-line to prepare a bitstream for transmission at some future point. For example, assuming that there may be considerable variation in the available bandwidth of a given transmission channel, the algorithm may be applied to a bitstream and re-ordered sequences may be calculated for several alternative bandwidths, such as 25%, 50% and 75%. In transmission, if problems are encountered, for example with pausing, the re-ordered sequence currently being used may be changed to one better suited to a lower available bandwidth. Equally, if transmission is proceeding better than expected, a sequence better suited for a higher bandwidth may be used. The invention may check the progress of transmission and/or playback a number of times during transmission and adjust the sequence accordingly. It is advantageous if this adjustment can be programmed dynamically whilst transmission is occurring. In this way, intermittent problems can de dealt with.

The part of the invention which re-orders a bitstream is located in an editor or in a server. The part of the invention which recovers the order is located in a client.

A further improvement is provided if the proportion of bits belonging to the primary frames is reduced by replacing some of the primary frames with B-frames. This reduces the compression efficiency slightly because B-frames are predicted from reference frames (P-frames) which are further away in time and so the B-frames and P-frames are less alike. This yields a poorer predicted B-frame and so more bits must be used to code the corresponding prediction error image. Since in this arrangement there are now more bits belonging to secondary frames which can be discarded, re-ordering the frames can provide even better protection against transmission delays and pauses.

One way of looking at the invention is to say that it sacrifices the safety times of the secondary frames in order to maintain maximum safety times for the primary frames.

The invention is particularly suitable for streaming-type multimedia retrieval.

While preferred embodiments of the invention have been shown and described, it will be understood that such embodiments are described by way of example only. For example, although the invention has been described applied to a temporally scalable bitstream, it can also be applied to other types of scalability, including spatial, spectral and SNR. Numerous variations, changes and substitutions will occur to those skilled in the art without departing from the scope of the present invention. Accordingly, it is intended that the following claims cover all such variations as fall within the spirit and scope of the invention.

Claims

Claims

1. A server for transmitting a data signal having a sequence of data units in a predetermined order over a transmission link the data units being sent in an order determined by their relative importance rather than their predetermined order.

2. A server according to claim 1 in which the data units represent a base layer and at least one enhancement layer.

3. A server according to claim 2 comprising re-ordering means to change the order of the data units.

4. A server according to any preceding claim in which the data signal is scalable.

5. A server according to claim 4 in which the signal is scalable in a domain selected from a group consisting of the temporal, the spatial, the spectral and the SNR domains.

6. A server according to any preceding claim which comprises an editor for providing the data signal.

7. A server according to any preceding claim in which the data signal represents a sequence of pictures to produce a moving image.

8. A server according to claim 7 in which the data signal represents a video sequence.

9. A server according to any preceding claim in which the data signal comprises multimedia data.

10. A data transmission system for transmitting a data signal having a sequence of data units in a predetermined order over a transmission link between a data

source and a data sink the data units being sent in an order determined by their relative importance rather than their predetermined order.

11. A transmission system according to claim 10 in which the data units each comprise a base layer and at least one enhancement layer and when re-ordered the base layer of a particular data unit has a greater safety time than of the or each enhancement layer of the particular data unit.

12. A transmission system according to claim 10 or claim 11 in which the source is a server.

13. A transmission system according to any of claims 10 to 12 in which the source is an editor.

14. A transmission system according to any of claims 10 to 13 in which the sink is a client.

15. A transmission system according to any of claims 10 to 14 in which the sink is a mobile terminal.

16. A transmission system according to claims 10 to 15 in which the sink is a mobile telephone.

17. A transmission system according to claims 10 to 16 in which means are provided to check the progress of transmission and to change the order being used to one better suited to available bandwidth.

18. A method of transmitting a data signal having a sequence of data units in a predetermined order over a transmission link between a data source and a data sink comprising the step of sending the data units in an order determined by their relative importance rather than their predetermined order.

19. A method according to claim 18 in which the data units are returned to their original sequence once they have been transmitted over the transmission link.

20. A method according to claim 18 or claim 19 in which the progress of transmission is checked and the order being used is changed to one better suited to available bandwidth.

21. A computer program product stored on a computer usable medium comprising computer readable program means for causing transmission of a data signal having a sequence of data units in a predetermined order over a transmission link the data units being sent in an order determined by their relative importance rather than their predetermined order.

22. A computer program product according to claim 21 comprising a server.

23. A computer program product according to claim 21 or claim 22 comprising an editor for providing a scalable data signal.

24. A computer program product according to any of claims 22 to 23 comprising re-ordering means for providing the layers of the or each data units with different safety times.

25. A data signal having a sequence of data units for transmission over a transmission link between a data source and a data sink the data units being in an order determined by their relative importance rather than their predetermined order.

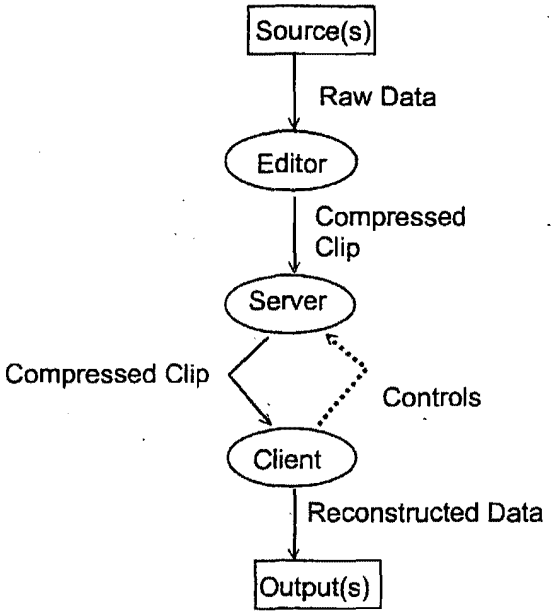

Source(s)
Raw Data
Editor
Compressed Clip
Server
Compressed Clip
Controls
Client
Reconstructed Data
Output(s)

Fig. 1

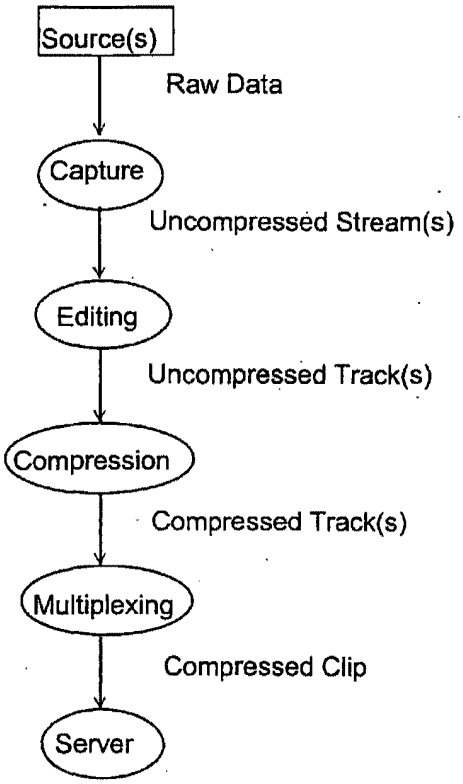

Source(s)
Raw Data
Capture
Uncompressed Stream(s)
Editing
Uncompressed Track(s)
Compression
Compressed Track(s)
Multiplexing
Compressed Clip
Server

Fig. 2

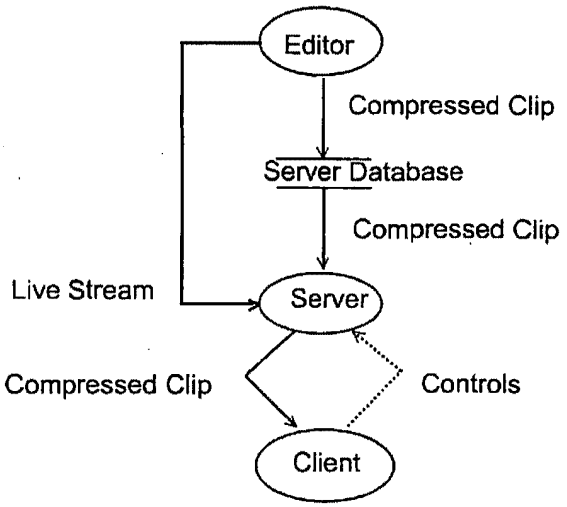

Editor
Compressed Clip
Server Database
Compressed Clip
Live Stream
Server
Compressed Clip
Controls
Client

Fig. 3

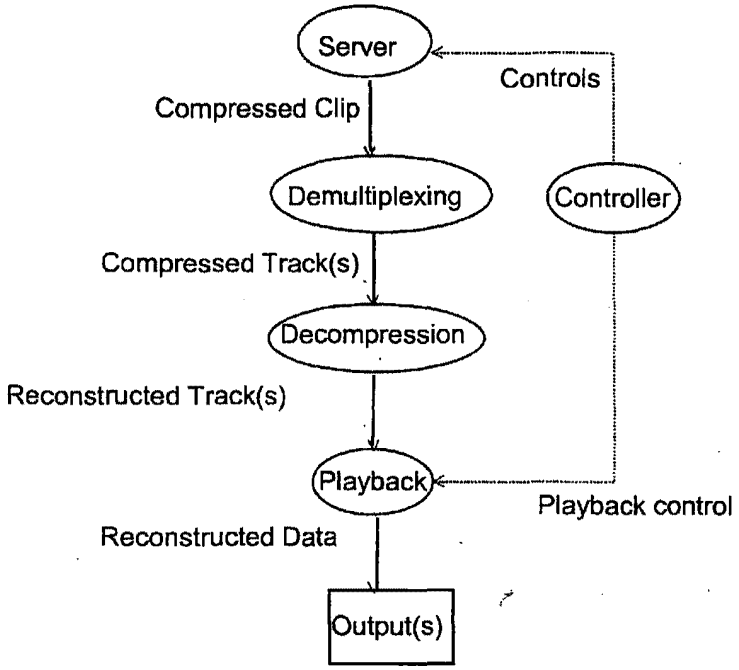

Server
Controls
Compressed Clip
Demultiplexing
Controller
Compressed Track(s)
Decompression
Reconstructed Track(s)
Playback
Playback control
Reconstructed Data
Output(s)

Fig. 4

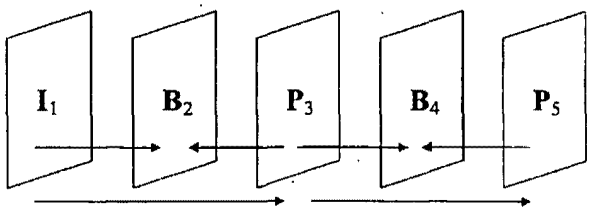


Fig. 5

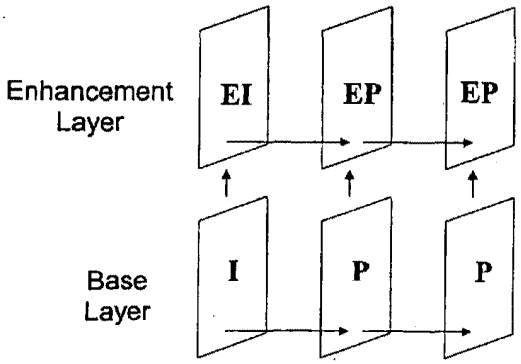


Fig. 6

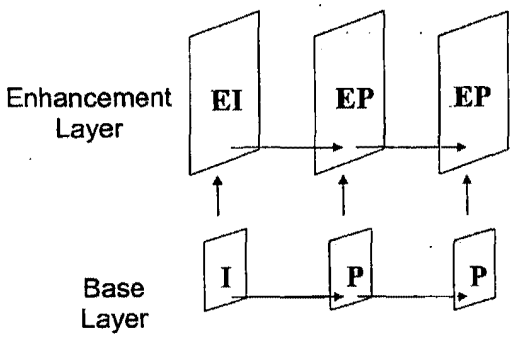

Enhancement Layer
EI
EP
EP
Base Layer
I
P
P

Fig. 7

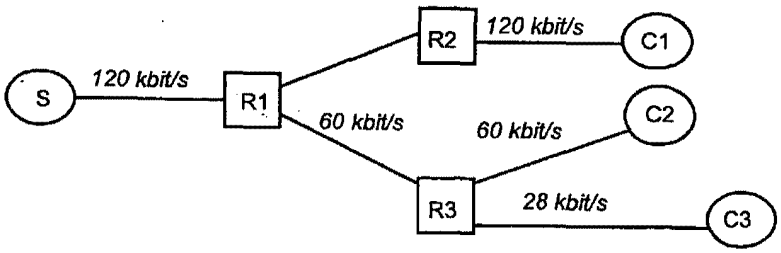

S
120 kbit/s
R1
60 kbit/s
R2
120 kbit/s
C1
60 kbit/s
C2
R3
28 kbit/s
C3

Fig. 8

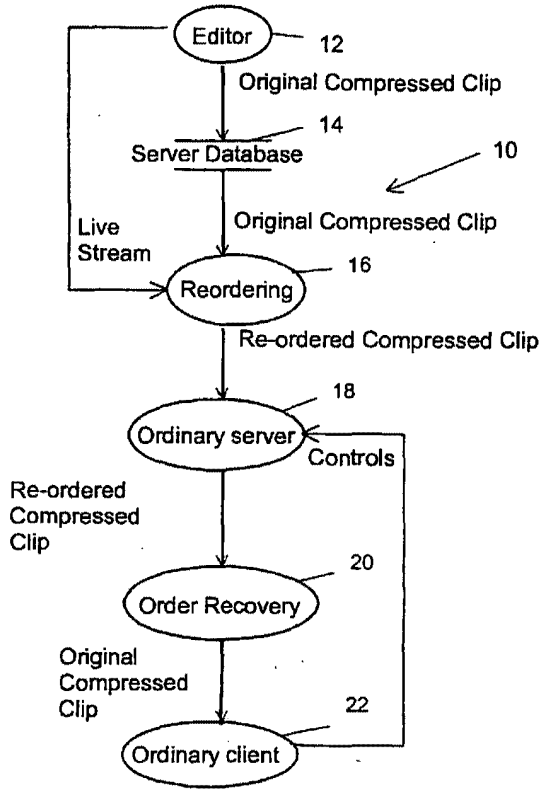

Editor
12
Original Compressed Clip
14
Server Database
10
Original Compressed Clip
Live
Stream
16
Reordering
Re-ordered Compressed Clip
18
Ordinary server
Controls
Re-ordered
Compressed
Clip
20
Order Recovery
Original
Compressed
Clip
22
Ordinary client

Fig. 9

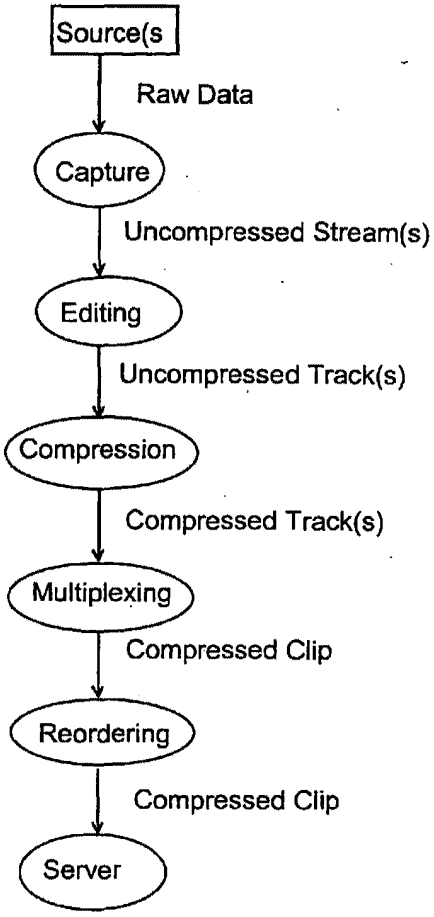

Source(s
Raw Data
Capture
Uncompressed Stream(s)
Editing
Uncompressed Track(s)
Compression
Compressed Track(s)
Multiplexing
Compressed Clip
Reordering
Compressed Clip
Server

Fig. 10

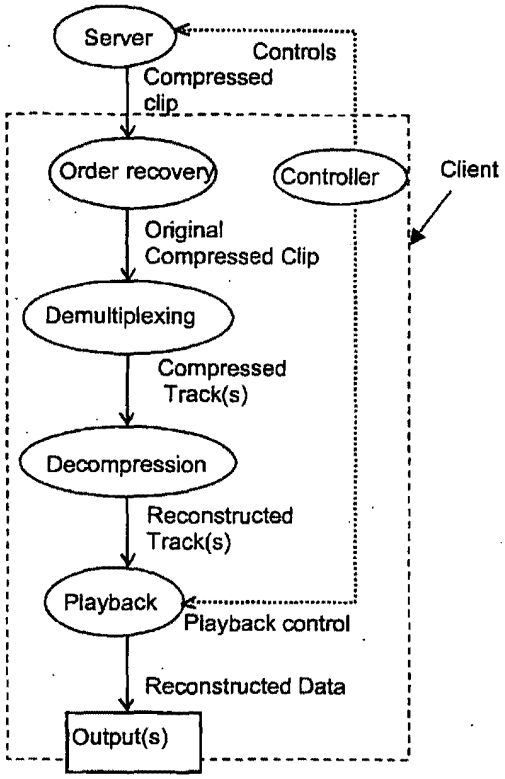

Server
Controls
Compressed clip
Order recovery
Controller
Client
Original Compressed Clip
Demultiplexing
Compressed Track(s)
Decompression
Reconstructed Track(s)
Playback
Playback control
Reconstructed Data
Output(s)

Fig. 11

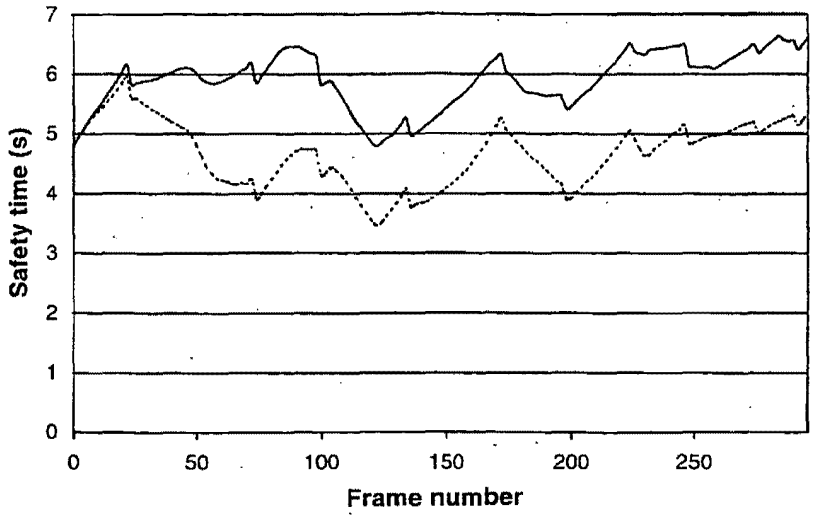


Fig. 12

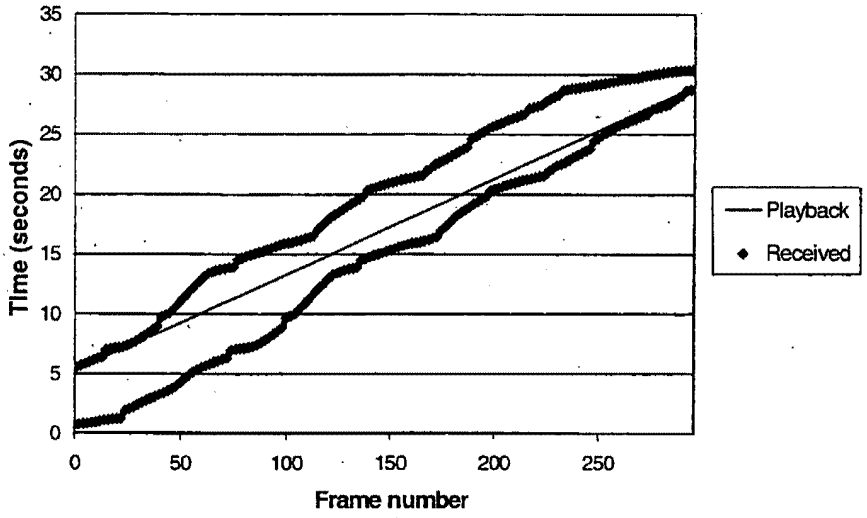

35
30
25
20
15
10
5
0
Time (seconds)
0
50
100
150
200
250
Frame number
Playback
Received

Fig. 13

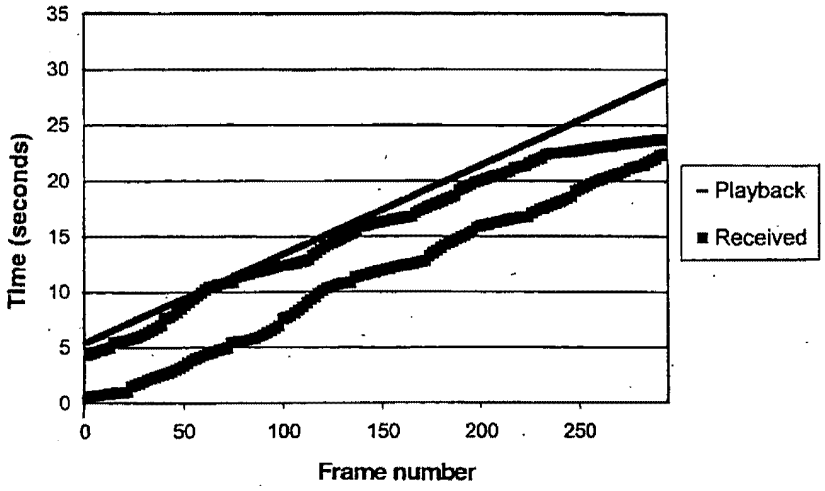

35
30
25
20
15
10
5
0
Time (seconds)
0
50
100
150
200
250
Frame number
Playback
Received

Fig. 14

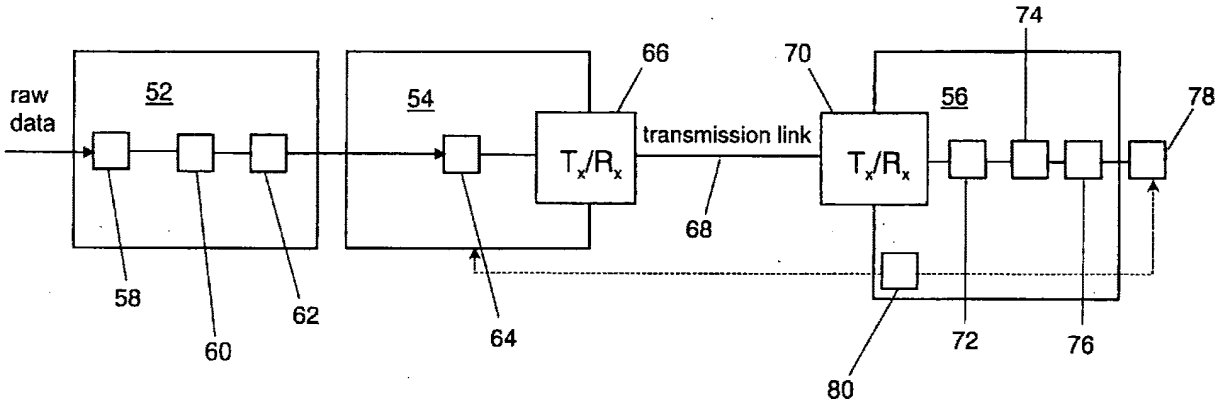

raw
data
52
54
66
70
56
74
78
transmission link
T_x/R_x
T_x/R_x
68
58
60
62
64
72
76
80

Fig. 15

INTERNATIONAL SEARCH REPORT

Interna al Application No
PCT/FI 00/00363

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 H04N5/00 H04N7/26

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 H04N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
WPI Data, PAJ, EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	PATENT ABSTRACTS OF JAPAN vol. 018, no. 566 (E-1622), 28 October 1994 (1994-10-28) & JP 06 205384 A (NEC CORP), 22 July 1994 (1994-07-22) abstract	1,6-10, 12-23,25
Y		2-5,11, 24

Y	US 5 564 001 A (LEWIS SCOTT W) 8 October 1996 (1996-10-08) column 2, line 58 -column 3, line 30 column 4, line 58 - line 65 column 7, line 4 - line 14 column 7, line 43 - line 47 abstract	2-5,11, 24
A		1,6-10, 12-23,25

	-/--	

[X] Further documents are listed in the continuation of box C. [X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
9 August 2000	- 9. 10. 2000
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	**Authorized officer** Jesper Bergstrand

Form PCT/ISA/210 (second sheet) (July 1992)

page 1 of 2

INTERNATIONAL SEARCH REPORT

Interm al Application No
PCT/FI 00/00363

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	EP 0 634 871 A (AT & T CORP) 18 January 1995 (1995-01-18) the whole document -----	1-25

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
28 June 2001 (28.06.2001)

PCT

(10) International Publication Number
WO 01/47283 A1

(51) International Patent Classification[7]: H04N 7/26

(21) International Application Number: PCT/US00/09584

(22) International Filing Date: 10 April 2000 (10.04.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/171,774 22 December 1999 (22.12.1999) US

(71) Applicant: GENERAL INSTRUMENT CORPORATION [US/US]; 101 Tournament Drive, Horsham, PA 19044 (US).

(72) Inventors: WAN, Wade, K.; Apartment 501, 12-16 Ellery Street, Cambridge, MA 02138 (US). CHEN, Xuemin; 8560 Foxcroft Place, San Diego, CA 92129 (US). LUTHRA, Ajay; 13695 Winstanley Way, San Diego, CA 92130 (US).

(74) Agent: LIPSITZ, Barry, R.; Building No. 8, 755 Main Street, Monroe, CT 06468 (US).

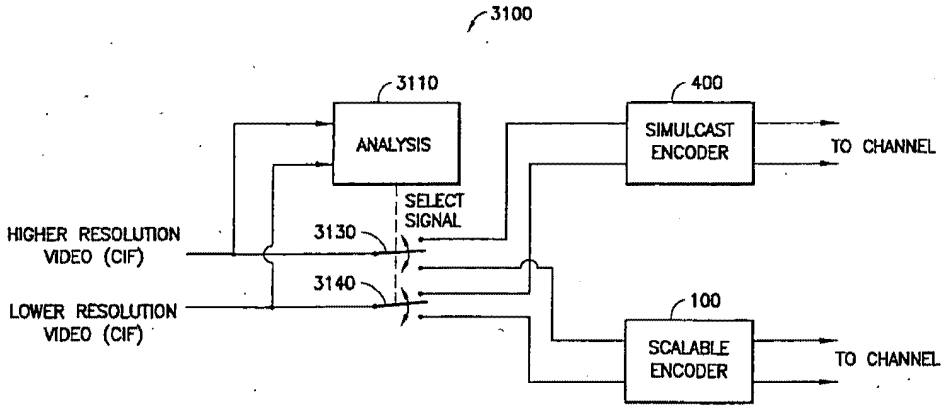
(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *With international search report.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: VIDEO COMPRESSION FOR MULTICAST ENVIRONMENTS USING SPATIAL SCALABILITY AND SIMULCAST CODING



(57) Abstract: An optimal system for determining whether simulcast coding (400) or spatial scalability coding (100) should be used to encode video for clients with a specific communication link. Operating points (A', B', C') for both simulcast coding and spatial scalability are also determined. Adaptive switching (3130, 3140) is provided, with the operating points and decision boundaries being used to guide the switching to optimize the quality of the higher-resolution data based on a bit allocation to the associated lower-resolution data. A system for determining the point (A', C') of equal quality in both layers of simulcast and spatial scalability coding is also disclosed. The proportion of bits allocated to the base layer to achieve equal quality is essentially independent of the total bit rate for both simulcast and spatial scalability.

WO 01/47283 A1

VIDEO COMPRESSION FOR MULTICAST ENVIRONMENTS USING SPATIAL SCALABILITY AND SIMULCAST CODING

BACKGROUND OF THE INVENTION

This application claims the benefit of U.S. Provisional Application No. 60/171,774, filed December 22, 1999.

The present invention relates to digital television and the like, and more particularly to a video coding scheme for multicast applications. The invention is particularly suitable for providing a streaming video server for multicast video over computer networks, such as Internet protocol (IP) networks. A multicast transmission can use simulcast ("Sim.") or spatial scalability ("SS") coding.

Usually, three major factors in a multicast video service need to be considered:

(1) The costs of reaching the audience (from the video-coding point of view);

(2) Quality of Service (e.g. visual quality); and

(3) Encoding complexity vs. decoding complexity.

Some multicast systems have chosen to use simulcast coding (the independent coding of bitstreams to achieve video scalability) exclusively. The simulcast approach does not require additional encoder or decoder complexity and thus satisfies the third factor listed above.

For some application environments, codec complexity may not be an issue and the first two factors are the main concerns. For these services, the selection of the video compression (coding) scheme to be used often follows the rule that for a given total combined allocation of two or more service rates, a coding scheme that can achieve the highest peak signal-to-noise ratio (PSNR) for clients is desired. Also, for a given quality requirement (e.g., PSNR for each bitstream), a coding scheme that can achieve the smallest total bit allocation is desired.

Intuitively, scalable bitstreams (dependently coded bitstreams) are expected to perform better than simulcast coding. This reasoning suggests that a multicast implementation without complexity constraints should only use spatial scalability. However, the present inventors have found that simulcast coding outperforms spatial scalability for certain operating regions. This surprising discovery enables a determination to be made as to which approach to use for a given application and to provide an adaptive switching technique between the two coding approaches.

It would be advantageous to provide an optimal method for selecting between spatial scalability and simulcast coding for multicast video services. Such a method should enable a determination to be made as to whether simulcast coding or spatial scalability should be used to encode video for clients with a specific

communication link. It would be further advantageous if, in addition to guiding decisions for specific cases, the methodology could be used to construct decision regions to guide more general scenarios, or used to adaptively switch between the two approaches. Operating points for both simulcast coding and spatial scalability, in terms of bit allocations among clients, should also be determinable by using such a method.

The present invention provides a system having the aforementioned and other advantages.

SUMMARY OF THE INVENTION

An optimal technique is provided for selecting between spatial scalability and simulcast coding to provide an efficient compression algorithm. In particular, simulcast coding can outperform spatial scalability when a small proportion of bits is allocated to the base layer.

A technique is also provided for determining whether simulcast coding or spatial scalability should be used to encode video for clients with a specific communication link. Operating points for both simulcast coding and spatial scalability are also determined. Adaptive switching between the two approaches is also provided, with the operating regions being used to guide the switching.

The invention also provides a method for determining the point of equal quality in both layers of simulcast coding. The proportion of bits allocated to the base layer to achieve equal quality is independent of the total bit rate for both simulcast and spatial scalability.

Corresponding methods and apparatuses are presented.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a block diagram of a general scalable coder/decoder (CODEC) with two layers;

FIG. 2 illustrates temporal scalability with two layers;

FIG. 3 illustrates spatial scalability with two layers;

FIG. 4 is a block diagram of a system for providing simulcast coding with two bitstreams;

FIG. 5 is a plot of PSNR vs. bit rate for the single layer coding of the QCIF carphone video sequence;

FIG. 6 is a plot of PSNR vs. bit rate for the single layer coding of the CIF carphone video sequence;

FIG. 7 is a plot of PSNR vs. the fraction of total bits allocated to the lower-resolution stream, for QCIF/CIF simulcast of the carphone video sequence for total bit rates of 0.29, 0.32, and 0.35 Mbps;

FIG. 8 is a plot illustrating an example of the iterations needed to obtain the point of equal quality in both layers of simulcast coding for the QCIF/CIF carphone video sequence.

FIG. 9 is a plot of PSNR vs. total bit rate for a QCIF/CIF simulcast, where the lower-resolution stream and higher-resolution stream have the same PSNR for a given total bitrate;

FIG. 10 is a plot of the fraction of total bits

allocated to the lower-resolution stream vs. the total bit rate for a QCIF/CIF simulcast, where the lower-resolution stream and higher-resolution stream have the same PSNR for a given total bitrate;

FIG. 11 is a plot of PSNR vs. total bit rate for QCIF/CIF spatial scalable coding of the carphone video sequence;

FIG. 12 is a plot of PSNR vs. the fraction of total bits allocated to the base layer, for QCIF/CIF spatial scalable coding of the carphone video sequence for total bit rates of 0.29, 0.32 and 0.35 Mbps;

FIG. 13 is a plot of PSNR vs. total bit rate for QCIF/CIF spatial scalability, where the base layer and the enhancement layer for a given video sequence have the same PSNR for a given total bitrate;

FIG. 14 is a plot of the fraction of total bits allocated to the base layer vs. total bit rate for QCIF/CIF spatial scalability, where the base layer and the enhancement layer of a given video sequence have the same PSNR for a given total bitrate;

FIG. 15 is a plot of PSNR vs. total bit rate for QCIF/CIF simulcast and spatial scalable coding of the carphone video sequence, where the QCIF and CIF video sequences have the same PSNR for a given total bitrate;

FIG. 16 is a plot of PSNR vs. total bit rate for QCIF/CIF simulcast and spatial scalable coding of the news video sequence, where the QCIF and CIF video sequences have the same PSNR for a given total bitrate;

FIG. 17 is a plot of PSNR vs. total bit rate for QCIF/CIF simulcast and spatial scalable coding of the foreman video sequence, where the QCIF and CIF video sequences have the same PSNR for a given total bitrate;

FIG. 18 is a plot of PSNR vs. total bit rate for QCIF/CIF simulcast and spatial scalable coding of the basket video sequence, where the QCIF and CIF video sequences have the same PSNR for a given total bitrate;

FIG. 19 is a plot of PSNR vs. total bit rate for QCIF/CIF simulcast and spatial scalable coding of the silentvoice video sequence, where the QCIF and CIF video sequences have the same PSNR for a given total bitrate;

FIG. 20 is a plot of PSNR vs. total bit rate for QCIF/CIF simulcast and spatial scalable coding of the bus video sequence, where the QCIF and CIF video sequences have the same PSNR for a given total bitrate;

FIG. 21 is a plot of PSNR vs. the fraction of total bits allocated to the base layer (for SS) or to the lower-resolution stream (for Sim.), for the carphone video sequence and a total bandwidth of 0.29 Mbps, which also illustrates the tradeoff between base and enhancement layers (for SS), and between the lower and higher resolution streams (for Sim.);

FIG. 22 is a plot of PSNR vs. the fraction of total bits allocated to the base layer (for SS) or to the lower-resolution stream (for Sim.), for the carphone video sequence and a total bandwidth of 0.32

Mbps;

FIG. 23 is a plot of PSNR vs. the fraction of total bits allocated to the base layer (for SS) or to the lower-resolution stream (for Sim.), for the carphone video sequence and a total bandwidth of 0.35 Mbps;

FIG. 24 is a plot of PSNR vs. the bit rate for the enhancement layer (for SS) or the higher-resolution stream (for Sim.) for the carphone video sequence and a fixed bit rate of 0.29 Mbps for the base layer (for SS) or the lower-resolution stream (for Sim.);

FIG. 25 is a plot of PSNR vs. the bit rate for the enhancement layer (for SS) or the higher-resolution stream (for Sim.), for the carphone video sequence and a fixed bit rate of 0.05 Mbps for the base layer (for SS) or the lower-resolution stream (for Sim.);

FIG. 26 is a plot of the bit rate for the enhancement layer (for SS) or the higher-resolution stream (for Sim.) vs. the bit rate for the base layer (for SS) or the lower-resolution stream (for Sim.), for simulcast and spatial scalability decision boundaries;

FIG. 27 is a plot of normalized bit rate for the enhancement layer (for SS) or the higher-resolution stream (for Sim.) vs. normalized bit rate for the base layer (for SS) or the lower-resolution stream (for Sim.), for simulcast and spatial scalability decision boundaries;

FIG. 28 is a block diagram illustrating an example

of multicast broadcasting using simulcast coding;

FIG. 29 is a block diagram illustrating an example of multicast broadcasting using spatial scalability coding;

FIG. 30 is a plot of bit rate for the enhancement layer (for SS) or the higher-resolution stream (for Sim.) vs. bit rate for the base layer (for SS) or the lower-resolution stream (for Sim.), for simulcast and spatial scalability decision regions; and

FIG. 31 illustrates an adaptive simulcast/spatial scalability encoder apparatus in accordance with the present invention.

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides techniques for selecting between simulcast coding and spatial scalability for multicast services, such as multicast video.

A brief review of general scalable coding (temporal scalability in addition to spatial scalability) is first provided. Simulations have been performed with an MPEG-4 codec to gain insight into the issues involved in transmitting the same video sequence at multiple spatial resolutions. In addition to deciding between simulcast coding and spatial scalability, one issue is the proper allocation of bits to the different layers.

General guidelines are provided and a scenario for achieving equal quality in both layers is examined in detail. The results obtained using quarter common intermediate format (QCIF) and common intermediate format (CIF) resolution sequences may be directly applied to applications such as video transmission in a multicast environment.

While QCIF is used as an example of lower-resolution data, and CIF is used as an example of higher-resolution data, these are examples only, and other suitable data formats can be used.

CIF was developed so that computerized video images can be shared from one computer to another. An

image that is digitized to CIF has a resolution of 352 x 288 or 352 x 240, which is essentially one-half of the resolution of CCIR 601. The CCIR 601 recommendation of the International Radio Consultative Committee for the digitization of color video signals deals with color space conversion from RGB to YCrCb, the digital filters used for limiting the bandwidth, the sample rate (defined as 13.5 MHz), and the horizontal resolution (720 active pixels).

Many applications desire the capability to transmit and receive video at a variety of resolutions and/or qualities. One method to achieve this is with scalable or layered coding, which is the process of encoding video into an independent base layer and one or more dependent enhancement layers. This allows some decoders to decode the base layer to receive basic video and other decoders to decode enhancement layers in addition to the base layer to achieve higher temporal resolution, spatial resolution, and/or video quality.

The general concept of scalability is illustrated in FIG. 1 for a codec with two layers. Note that additional layers can be used. The scalable encoder 100 takes two input sequences and generates two bitstreams for multiplexing at a mux 140. Specifically, the input base video stream or layer is processed at a base layer encoder 110, and upsampled at a midprocessor 120 to provide a reference image for

predictive coding of the input enhanced video stream or layer at an enhancement layer encoder 130.

Note that coding and decoding of the base layer operate exactly as in the non-scalable, single layer case. In addition to the input enhanced video, the enhancement layer encoder uses information about the base layer provided by the midprocessor to efficiently code the enhancement layer. After communication across a channel, which can be, e.g., a computer network such as the Internet, or a broadband communication channel such as a cable television network, the total bitstream is demultiplexed at a demux 150, and the scalable decoder 160 simply inverts the operations of the scalable encoder 100 using a base layer decoder 170, a midprocessor 180, and an enhancement layer decoder 190.

The MPEG-2 standard defines scalable tools for spatial, temporal and quality (SNR) scalability. The main commercial applications that MPEG-2 was targeted to were digital video disks and digital television, applications where the additional functionality of scalability is often not used. Thus, there has been limited commercial interest in MPEG-2 scalable coding in the past. However, new applications such as streaming video could greatly benefit from scalability. One example where scalable coding may be useful is for video transmission in a multicast environment. Clients have a wide range of processing power, memory resources and available bandwidth. This requires a server to

provide different resolutions and/or qualities of video to be able to satisfy the different capabilities of their clients. The recently completed multimedia standard MPEG-4 version 1 offers two types of scalability: temporal and spatial. In addition to applying scalability to frames (pictures) of video, the standard also defines scalability for arbitrary shaped objects.

This document focuses on frame-based scalability, although the concepts of the invention are generally applicable to arbitrarily shaped objects. In addition to temporal and spatial scalability, a third type of scalable coding for quality scalability called Fine Granular Scalability (FGS) is currently being evaluated for inclusion in MPEG-4. A brief review of temporal and spatial scalability in MPEG-4 is presented before discussion of simulcast coding.

Temporal scalability permits an increase in the temporal resolution by using one or more enhancement layers in addition to the base layer.

FIG. 2 shows an example of temporal scalable coding with two layers. Basic video is obtained by decoding only the independent base layer 200, which is done in the same manner as in the non-scalable, single layer case. Use of the dependent enhancement layer 250 provides video with, e.g., seven times the temporal resolution of the basic video. The same spatial resolution is obtained whether or not the enhancement

layer 250 is used. A frame in the enhancement layer 250 can use motion compensated prediction from the previous or next frame in display order belonging to the base layer as well as the most recently decoded frame in the same layer.

Spatial scalability permits an increase in the spatial resolution by using enhancement layers in addition to the base layer. FIG. 3 shows an example of spatial scalable coding with two layers. Basic video is obtained by decoding only the independent base layer 300, which is done in the same manner as in the non-scalable, single layer case. Use of the dependent enhancement layer 350 provides video with, e.g., twice the spatial resolution of the basic video. The same temporal resolution is obtained whether or not the enhancement layer is used. A frame in the enhancement layer can use motion compensated prediction from the temporally coincident frame in the base layer as well as the most recently decoded frame in the same layer.

Another method to transmit video at multiple resolutions or qualities is simulcast coding. FIG. 4 shows an example of simulcast coding with two bitstreams. For simulcast coding, the streams are independent, whereas scalable coding usually refers to an independent base layer with one or more dependently-coded enhancement layers. For comparison with scalable coding, one of the simulcast streams (termed a lower-resolution stream) has the same resolution as the base

layer, and the other simulcast bitstream (termed a higher-resolution stream) has the same resolution as the enhancement layer.

First and second input bitstreams are coded at corresponding video encoders 420 and 410, respectively. The input video #1 is assumed to be a lower-resolution stream, while the input video #2 is assumed to be a higher-resolution stream. This involves coding each representation independently and is usually less efficient than scalable coding since similar information in another bitstream is not exploited. The bitstreams are then multiplexed at a mux 430, transmitted across some channel, demultiplexed at a demux 440, and decoded independently at video decoders 470 and 460, respectively, in a simulcast decoder 450. Unlike scalable coding, no additional decoder complexity is required to decode the higher-resolution video. This may be important for commercial applications since additional decoder complexity often increases the cost of receivers.

This invention focuses on the performance of spatial scalability and its simulcast counterpart. Simulations have been performed with an MPEG-4 encoder on rectangular video to gain insight into the issues with transmitting video at a variety of spatial resolutions. One issue with layered coding is the proper allocation of bits between layers. In addition to examining the differences between simulcast coding

and spatial scalability, investigations in connection with the invention focus on determining guidelines for bit allocation.

Six different video sequences were examined, namely Basket, Bus, Carphone, Foreman, News, and Silentvoice. This set of sequences has a wide range of complexity, so the results should be generally applicable to other sequences. The Basket and Bus sequences have a large amount of motion and may stress most encoders. The News and Silentvoice sequences have large stationary backgrounds making them easy to compress efficiently. The original sequences were in CIF format (288 x 352 pixels) and QCIF format (144 x 176 pixels). The sequences were created by downsampling (without use of any anti-aliasing filter, i.e., decimation). Each sequence was 150 frames long and the source material and display frame rates were 30 frames per second. An MPEG-4 encoder was used to encode the simulcast and spatial scalable streams at various fixed quantization levels with no rate control. The parameters used for the simulations are shown in Table 1. "VOP" refers to a Video Object Plane, as known from the MPEG-4 standard.

	Lower Layer of Simulcast Streams	Upper Layer of Simulcast Streams	Base Layer of Scalable Streams	Enhancement Layer of Scalable Streams
Q_I	{4, 6, 8, …, 24}	{4, 6, 8, …, 24}	{4, 6, 8, …, 24}	{6, 8, 10, …, 24}
Q_P	min (1.4Q_I , 32)	min (1.4Q_I , 32)	min (1.4Q_I , 32)	min (1.4Q_I , 32)
Q_B	min (1.8Q_I , 32)	min (1.8Q_I , 32)	min (1.8Q_I , 32)	min (1.8Q_I , 32)
Structure	IPPBPPB…	IPPBPPB…	IPPBPPB…	PBBPBBPBB…
M	3	3	3	12
N	12	12	12	---
Range	8	16	8	16

Q_I	Quantizer for I-VOPs
Q_P	Quantizer for P-VOPs
Q_B	Quantizer for B-VOPs
Structure	Picture Structure
M	Period between consecutive P-VOPs
N	Period between consecutive I-VOPs
Range	Search range for motion vectors

Table 1: Parameters of Simulations

The measure of quality here is the PSNR, which is defined to be the Peak Signal-to-Noise Ratio of the luminance (Y) component of the decoded video compared to the input video at the same resolution. While PSNR is the quality measure used here, other possible measures include, e.g., MSE (Mean Square Error) and a Perceptual Distortion Measure. PSNR as used here is defined to be 20*log10(MSE) (dB).

Single Layer Coding

Single layer coding results for the QCIF and CIF Carphone sequences are shown in FIG. 5 and FIG. 6. The legend "150-30fps" indicates a sequence of 150 frames at 30 frames per second. Note that 150 frames was an arbitrarily chosen length and many suitable sequence lengths could be used. Conceptually, switching between spatial scalability and simulcast coding can occur as frequently as at every picture. However, this may result in syntax problems. Switching between groups of pictures (GOPs) is a realistic possibility.

Moreover, here and in the other figures, the bit rate is an average bit rate over a sequence. The circles denote the empirical results and the dotted lines 500, 600, respectively, represent logarithmic fits to the data using the following model:

$$\mathrm{PSNR} = A\ln(\mathrm{Bitrate}) + B .$$

The model allows each single layer to be represented by two parameters (constants) A and B (along with the range of bit rates where this model is valid). "ln" denotes the natural logarithm. "Bitrate" is the "x" parameter in the figures.

Simulcast Coding

A typical scenario encountered when transmitting multiple sequences is a constraint on the total bandwidth.

FIG. 7 is a plot of PSNR vs. fraction of total bits allocated to the lower-resolution stream, for a

total bandwidth of 0.29, 0.32, and 0.35 Mbps for QCIF/CIF simulcast. The figure shows examples of the different qualities that can be achieved by limiting the total bandwidth to 0.29, 0.32 and 0.35 Mbps, respectively, for simulcast ("Sim.") transmission of the Carphone QCIF and CIF sequences. Both qualities are plotted as functions of the fraction of total bits allocated to the lower-resolution stream (i.e., the QCIF stream). In particular, the solid lines 700, 710, 720 represent the PSNR of the lower-resolution (QCIF) sequence for total bit rates of 0.29, 0.32, and 0.35 Mbps, respectively. The dotted lines 750, 760, 770 represent the PSNR of the higher-resolution (CIF) sequence for total bit rates of 0.29, 0.32, and 0.35 Mbps, respectively.

Note the monotonicity of the data for both layers. That is, the PSNR either increases or decreases steadily (without a peak or valley).

While some of the figures refer to the fraction of bits that are allocation to the base layer or enhancement layer (for SS), or to the lower-resolution stream or higher-resolution stream (for Sim.), note that these values can also be expressed in terms of an absolute number of bits, a percentage, a fraction or percentage of a fixed reference value, or any other linear or non-linear metric or scale.

The independence of the single streams in simulcast coding causes an improvement in quality with

an allocation of more bits to the QCIF stream, and degradation of the other, (CIF) stream. This makes sense since the additional bits that are given to one stream are taken away from the other stream to maintain the same total bandwidth.

The curve crosses in the figures denote the point where the functions intersect, i.e., the point where the PSNR of both streams is equal. This point can be found using the logarithmic fits to the data of each stream and a bisection algorithm since the PSNR of each stream is a monotonic function of the bit rate of the lower-resolution stream.

In particular, the monotonicity of the PSNR for both layers implies that if the functions intersect, they will intersect at only one point. In some cases, there may be no intersection point due to insufficient or excess total bandwidth. An additional exit condition checking the difference in bit rates can be used to determine if there is no intersection. This has been omitted from the following algorithm for simplicity. Therefore, the algorithm below assumes that a proper total bit rate has been selected allowing the functions to intersect. A bisection algorithm using the difference in PSNR between the layers can be used to find the point of equal quality. An example algorithm follows:

1. Assume fixed total bitrate R_{TOT} and a threshold for convergence $T > 0$.

2. Let R_1 and R_2 be the minimum and maximum bitrates for single layer coding of the lower layer.

3. Let $R_3 = \frac{R_1 + R_2}{2}$.

4. Compute $PSNR_{lower}$ and $PSNR_{upper}$ for R_3 using the logarithmic fits to each layer:

$$PSNR_{lower} = A_{lower} \ln(R_3) + B_{lower}$$
$$PSNR_{upper} = A_{upper} \ln(R_{TOT} - R_3) + B_{upper}$$

where A_{lower} and B_{lower} are the parameters for the lower (lower-resolution) layer and A_{upper} and B_{upper} are the parameters for the upper (higher-resolution) layer.

5. Let DIFF = $PSNR_{lower}$ - $PSNR_{upper}$.

6. If the absolute value of DIFF is less than T,

the algorithm is finished and R_3 is the lower layer bitrate to achieve equal quality in both layers of simulcast coding.

Otherwise,

Set $R_1 = R_3$ if DIFF < 0.

Set $R_2 = R_3$ if DIFF > 0.

Go back to step 3.

This algorithm is just one example of how the crossover point can be found. Other techniques are possible, such as a linear interpolation.

FIG. 8 gives an example of the iterations needed to obtain the point 800 of equal quality in both streams of the QCIF/CIF simulcast coding of the

carphone sequence, for the 0.32 Mbps case. $R_1(i)$ and $R_2(i)$ represent the bit rates R_1 and R_2, respectively, for the lower-resolution stream at iteration i of the algorithm. Note that FIG. 8 shows the absolute bit rate of the lower-resolution stream, whereas FIG. 7 shows the fraction of the total bits that are allocated to the lower-resolution stream. It should be easily understood here and elsewhere how to convert between these two types of scales.

The point where the PSNRs intersect in FIG. 7 can be interpreted as the bit allocation where both streams are coded at approximately the same quality since PSNR is normalized with respect to picture size. Note that this point may not occur with some total bandwidths due to the limited dynamic range of each stream's coding.

The ability to transmit two streams of different resolutions that have roughly the same quality may be desirable in applications such as streaming video over the Internet with the resolutions used here. However, other applications may have different requirements. For example, consider two different agendas for the same simulcast system. One application may desire relatively higher quality in the lower-resolution stream to satisfy a larger number of receivers receiving the lower resolution. Another application may desire a relatively higher quality in the higher-resolution streams to satisfy the receivers receiving the higher resolution because of the higher cost of the

bandwidth required to receive the entire simulcast stream. The analysis in the following sections focuses on achieving the same quality for both streams, but this may not be the goal of some applications and this issue is revisited hereinbelow.

Additionally, FIG. 7 illustrates that the fraction of bits allocated to the lower-resolution stream to achieve equal quality in the two streams is essentially independent of the total bit rate. This is a very useful result and its significance can be seen by the following example. Assume that a multiplexer is combining simulcast bitstreams and has already determined the proper bit allocation between streams. The preceding result suggests that the multiplexer does not have to redetermine the proper bit allocation when reacting to a change in the total bandwidth. Instead, the proportion of bits allocated to each stream should remain the same.

FIGs 9 and 10 show the results of performing the same analysis as described above for different sequences at a wide variety of constrained total bandwidths. In particular, FIG. 9 shows the PSNR versus bitrate for the test sequences Basket 900, Bus 910, Carphone 920, Foreman 930, News 940, and Silentvoice 950.

The quality of both the lower-resolution (QCIF) and higher-resolution (CIF) streams is shown by only one curve for each sequence since the data in this plot

was obtained by requiring equal PSNR for both streams, and the bitrate shown is the total bitrate. For example, for the Basket sequence 900, when the total bitrate is 1 Mbps, both streams have a PSNR of approximately 24 dB. As the total bitrate is slowly increased, the PSNR of both streams increases, up to approximately 32 dB when the total bit rate is 4.5 Mbps.

FIG. 10 shows the fraction of total bits allocated to the lower-resolution streams versus total bitrate for the test sequences Basket 1000, Bus 1010, Carphone 1020, Foreman 1030, News 1040, and Silentvoice 1050. The data show that approximately 40 ± 4 % of the total bandwidth should be allocated to the lower-resolution stream to achieve equal quality in both streams of simulcast coding. This result can be used as a general guide for the bit allocation of simulcast bitstreams.

Spatial Scalable Coding

Results for the spatial scalable ("SS") coding of the Carphone QCIF and CIF sequences are shown in FIG. 11. The abscissa of each data point (on the horizontal axis) is the total bit rate of both streams, and the ordinate (on the vertical axis) represents the PSNR of the enhancement layer (the higher, CIF resolution) using spatial scalability. Each set of curves uses the same lower-resolution QCIF base layer (and therefore base layer bit rate and PSNR) with different enhancement layer bit rates. The circles denote

empirical results, and the dotted lines are logarithmic fits to each set of curves using the same base layer. Table 2 lists the bit rate (in Mbps) and PSNR of the QCIF base layer as well as the parameters (A and B) of the logarithmic fit ($Y = A \ln(X) + B$) for each set of curves.

Base Bitrate (Mbps)	Base PSNR	A	B
0.286	36.46	9.40	41.34
0.171	34.11	6.89	40.90
0.108	32.73	5.70	40.45
0.0835	31.69	5.36	40.26
0.0714	31.04	5.32	40.26
0.0625	30.46	5.36	40.34
0.0561	29.92	5.40	40.41
0.0518	29.48	5.47	40.49
0.0492	29.09	5.55	40.53
0.0472	28.76	5.62	40.61
0.0456	28.44	5.66	40.62

Table 2: Base (QCIF) Layer Characteristics of Spatial Scalable Coding for QCIF/CIF (Carphone)

In FIG. 11, the 0.286, 0.171, and 0.108 base layer bit rates are shown at curves 1100, 1110, and 1120, respectively. The remaining bit rates of 0.0835 through 0.0456 are shown at the aggregate curves 1130.

A constraint on the total bandwidth is examined for spatial scalable coding of the Carphone QCIF and CIF sequences. FIG. 12 is a plot of PSNR vs. the fraction of total bits allocated to the base layer, for

total bandwidths of 0.29, 0.32 and 0.35 Mbps for QCIF/CIF spatial scalability. The figure shows examples of the different qualities that can be achieved by limiting the total bandwidth to 0.29, 0.32 and 0.35 Mbps. The solid lines represent the PSNR of the base layer (QCIF) sequence and the dotted lines represent the PSNR of the enhancement layer (CIF) sequence. In particular, the solid lines 1200, 1210, 1220 represent the PSNR of the QCIF sequence for 0.29, 0.32, and 0.35 Mbps, respectively. The dotted lines 1250, 1260, 1270 represent the PSNR of the CIF sequence for 0.29, 0.32, and 0.35 Mbps, respectively.

Note the PSNR of the CIF sequences is not a monotonic function of the fraction of bits allocated to the base layer. That is, there is a peak in the PSNR for the CIF sequences near 20%.

As expected, the PSNR of the QCIF sequence is a monotonically increasing function of the fraction of total bits allocated to it. Moreover, it appears that allocating less than approximately 20% of the total bitstream to the base layer gives declining performance in both layers. The decline is relatively slight for the CIF layer, but rather sharp for the QCIF layer. This trend is also present with the other test sequences. This result is different from the one seen in the simulcast approach, where both layers are independent, and is due to the dependence of the CIF enhancement layer on the upsampled QCIF base layer.

This result of 20% allocation for the base layer can be a useful minimum boundary for the allocation of base layer bits for spatial scalable encoders.

Additionally, FIG. 12 shows that the fraction of bits allocated to the base layer to achieve equal quality in the two layers is essentially independent from the total bit rate. This result is similar to the conclusions obtained after analysis of the simulcast experiments and can be very useful for allocating spatial scalable bitstreams.

FIGs 13 and 14 show the results of performing the same analysis as described above for different sequences at a wide variety of constrained total bandwidths, where the base layer and enhancement layer have the same PSNR for a given total bit rate. In particular, FIG. 13 shows the PSNR versus total bitrate for the test sequences Basket 1300, Bus 1310, Carphone 1320, Foreman 1330, News 1340, and Silentvoice 1350. FIG. 14 shows the fraction of total bits allocated to the base layer versus total bitrate for the test sequences Basket 1400, Bus 1410, Carphone 1420, Foreman 1430, News 1440, and Silentvoice 1450. The data of FIG. 14 shows that approximately 45 ± 5 % of the total bandwidth should be allocated to the base layer to achieve equal quality in both layers of spatial scalable coded bitstreams. This percentage can be used as a general guide for the bit allocation of spatial scalable bitstreams.

FIGs 15 through 20 show the simulcast and spatial scalability results for each test sequence, where the QCIF and CIF video sequences have the same PSNR for a given total bit rate.

In particular, FIG. 15 shows the PSNR for spatial scalability 1500 and simulcast 1510 for the carphone sequence, FIG. 16 shows the PSNR for spatial scalability 1600 and simulcast 1610 for the news sequence, FIG. 17 shows the PSNR for spatial scalability 1700 and simulcast 1710 for the foreman sequence, FIG. 18 shows the PSNR for spatial scalability 1800 and simulcast 1810 for the basket sequence, FIG. 19 shows the PSNR for spatial scalability 1900 and simulcast 1910 for the silentvoice sequence, and FIG. 20 shows the PSNR for spatial scalability 2000 and simulcast 2010 for the bus sequence.

The operating regions of simulcast coding and spatial scalability are often different, with spatial scalability being the only option at relatively low total bit rates, and simulcast coding at relatively high total bit rates. Specifically, at relatively low bit rates, spatial scalability can be used if a coarse quantizer is used for residual coding. Simulcast coding may not be possible since the bandwidth may be too low to encode sequences at the higher resolution even with the coarsest quantizer. Note that much more information must be encoded for the simulcast case

since no information is available from the base layer. That is, there usually is a lot less signal energy in the residual (the difference between the uncoded enhancement layer and an upsampled version of the decoded base layer) than in the original signal. The enhancement layer of spatial scalability can be thought of as encoding the residual while the second, higher-resolution stream of simulcast coding is encoding the original, high resolution signal. Since we are using the same range of quantizers for both the enhancement layer of spatial scalability and the second layer of simulcast coding, it is not surprising that there are different ranges for the coded bitrates for the two methods.

Except for some regions with the Basket sequence, there is an improvement in quality gained by using spatial scalability at bit rates where both simulcast coding and spatial scalability are possible. Table 3 lists the range of PSNR improvements for each sequence using spatial scalability where both simulcast and scalable coding are possible. The negative value for the minimum PSNR improvement for the Basket sequence indicates that simulcast coding achieves higher quality video for part of the common operating region.

Note that the decision between simulcast coding and spatial scalability for a commercial application generally involves more than looking at the differences in PSNR or other quality measure. The lower layer bit

rate is smaller for simulcast coding, which favors simulcast coding since less bandwidth is required for reception of only the base layer. Additional decoder complexity is also required to decode spatial scalability bitstreams.

Sequence	Maximum PSNR Improvement Using Spatial Scalability	Minimum PSNR Improvement Using Spatial Scalability
Carphone	0.74	0.41
News	1.06	0.79
Foreman	0.75	0.68
Basket	0.71	-0.25
Silentvoice	1.27	1.18
Bus	0.65	0.11

Table 3: Range of PSNR Improvements Using Spatial Scalability Where Both Simulcast and Scalable Coding Are Possible

The previous analysis focused on achieving equal PSNR in both layers. As discussed earlier, this may not be the goal of some applications. A different view of this operating point reveals some additional insight into the general differences between simulcast coding and spatial scalability.

FIGs 21 through 23 show the results of both simulcast and scalable coding for the QCIF and CIF Carphone sequences with fixed total bandwidths of 0.29, 0.32 and 0.35 Mbps, respectively. The results are plotted as functions of the fraction of bits allocated

to the base layer (for SS), or the lower-resolution stream (for Sim.).

In particular, FIG. 21 shows the PSNR for 0.29 Mbps for simulcast CIF 2100, spatial scalability CIF 2110, and QCIF 2120 (which is the same for simulcast or spatial scalability). FIG. 22 shows the PSNR for 0.32 Mbps for simulcast CIF 2200, spatial scalability CIF 2210, and QCIF 2220. FIG. 23 shows the PSNR for 0.35 Mbps for simulcast CIF 2300, spatial scalability CIF 2310, and QCIF 2320.

Note that simulcast ("Sim.") outperforms spatial scalability ("SS") if a relatively small percentage of the total bit rate is assigned to the base layer. One general trend appears to be the increasing advantage of spatial scalability with more bits allocated to the base layer.

As an example, FIG. 21 provide visual markers to aid the following discussion. Point A' is the operating point for equal PSNR in both streams using simulcast coding. Point B' is the corresponding point for spatial scalability using the same amount of bits allocated to the base layer as Point A'. Point C' is the operating point for equal PSNR in both layers using spatial scalability. Note that the use of the operating points that achieves equal quality in both streams or layers (Points A' and C', respectively) causes different amounts of bits to be allocated to the lower-resolution stream or base layer depending,

respectively, on whether simulcast (Point A') or spatial scalability (Point C') is used.

Comparison of simulcast coding and spatial scalability with the same amount of bits allocated to the lower-resolution stream and base layer (Points A' and B', respectively) shows that spatial scalability results in a higher PSNR at the higher CIF resolution. Note that the PSNR at the lower QCIF resolution is the same for both simulcast coding and spatial scalability. Moreover, this is different from the gain in both the QCIF and CIF resolutions obtained by using the operating point for equal quality in both layers (Point C'). One way to interpret this concept is that part of the PSNR gain in the enhancement layer by using spatial scalable coding can be "exchanged" for an increase in the base layer by "moving" bits from the enhancement to the base layer. In fact, more bits can also be "moved" from the base layer to the enhancement layer. This concept can be visualized by simultaneously moving along the QCIF and CIF curves in FIG. 21. This allows different distributions, such as the points between Points B' and C' (for SS), or between Points A' and C' (for Sim.), to be achievable.

In general, the bit allocation problem involves two bit rates (the base and enhancement bitstreams for SS., and the lower-resolution and higher-resolution bitsreams for Sim.), and the choice between simulcast coding and spatial scalability. Note that more than

two streams or layers may be used, in which case the bit allocation problem is extended accordingly.

One method to obtain this three-dimensional data (assuming two streams or layers) is to fix the rate of the base layer (or lower-resolution stream), and decide between simulcast coding and spatial scalability for different bit rates for the enhancement layer or higher-resolution stream. By combining the data at different base layer (lower-resolution stream) rates, the complete three-dimensional data can be constructed.

FIGs 24 and 25 are examples of fixing the bit rate (and therefore, PSNR) of the base layer (or lower-resolution stream), for the transmission of the Carphone QCIF and CIF sequences. In particular, FIG. 24 shows the PSNR of the higher-resolution stream or enhancement layer, respectively, for simulcast 2400 and spatial scalability 2410 at 0.29 Mbps, and FIG. 25 shows the corresponding PSNR for simulcast 2500 and spatial scalability 2510 at 0.05 Mbps. Note that the curve for spatial scalability has a smaller dynamic range.

In FIG. 24, fixing the QCIF data results in a PSNR of 36.45 for that data. The total bit rate is then 0.29 Mbps + the enhancement layer or higher-resolution stream layer bit rate. The data point at (0 Mbps, 30 dB) results if no enhancement layer data is used, i.e., the base layer is simply upsampled (using bilinear interpolation) to create the enhancement layer.

In FIG. 25, the fixed QCIF bit rate results in a PSNR of 28.36 for that data. The total bit rate is then 0.05 Mbps + the enhancement layer or higher-resolution stream bit rate. The data point at (0 Mbps, 27 dB) results if no enhancement layer data is used, i.e., the base layer is simply upsampled (using bilinear interpolation) to create the enhancement layer. The general trend is for spatial scalability to be more efficient at lower enhancement layer/higher-resolution stream bit rates, while simulcast coding is more efficient at higher enhancement layer/higher-resolution stream bit rates. An important result is to determine the boundary where simulcast coding and spatial scalability are equivalent. This boundary can then be used to determine whether one should use simulcast coding or spatial scalability.

The functions for simulcast coding and spatial scalability may not intersect, but the curves can be extrapolated to find an intersection point by fitting both curves to logarithms and finding the intersection of the logarithmic fits. In particular, assume the two logarithmic functions are:

$$Y_1 = A\ln\left(X\right) + B$$

$$Y_2 = C\ln\left(X\right) + D$$

Y1 is the PSNR for one curve, e.g., the simulcast coding curve, with curve fit constants A and B. Y2 is the PSNR for the other curve, e.g., the spatial

scalability curve, with curve fit constants C and D. Solving for the point of intersection between the two logarithmic functions yields:

$$Y_1 = Y_2$$
$$A\ln(X) + B = C\ln(X) + D$$
$$(A - C)\ln(X) = D - B$$
$$\ln(X) = \frac{D - B}{A - C}$$
$$X = e^{\frac{D-B}{A-C}}$$

The crosses in the figures (point 2420 in FIG. 24, and point 2520 in FIG. 26) represent the estimated points of intersection.

FIG. 26 shows the results of applying this technique of estimating the boundary between simulcast coding and spatial scalability for all the test sequences. This figure shows the test sequences Basket 2600, Bus 2610, Carphone 2620, Foreman 2630, News 2640, and Silentvoice 2650. For each video sequence, the data indicates that operating points above the curve should use simulcast coding and points below the curve should use spatial scalability to obtain the highest PSNR for the CIF video sequence (e.g., the enhancement layer or higher-resolution stream). Note that specifying the base bit rate uniquely determines the PSNR of the QCIF sequence since this resolution is single layer coded. The decision boundaries tend to have the same shape with different scales.

FIG. 27 demonstrates the similarity of the decision boundaries this by normalizing the base layer and the enhancement layer bit rates (for SS), and by normalizing the lower-resolution stream and higher-resolution stream bit rates (for Sim.), using the range of the data and the following formula:

$$Normalized\ Bitrate = \frac{Bitrate - Minimum}{Maximum - Minimum}.$$

This formula maps the minimum absolute bit rate to a zero normalized bit rate and the maximum absolute bit rate to a normalized bit rate of one.

FIG. 27 shows normalized decision boundaries for the test sequences Basket 2700, Bus 2710, Carphone 2720, Foreman 2730, News 2740, and Silentvoice 2750.

The results above allow one to choose between simulcast and spatial scalability strictly on the basis of which mode provides higher quality. Considering other issues such as the additional receiver complexity required for spatial scalability may require not only determining which mode is better, but how much improvement is obtained, especially when factors support the other mode. This requires looking at the three-dimensional data. Note that determining which mode is better regardless of the PSNR difference is equivalent to a projection of the three-dimensional data onto a two-dimensional space. Visualizing and establishing decision criteria for surfaces is

difficult for general scenarios and may be best handled in a case-by-case manner.

Multicast environment

The present invention can be applied to video broadcasting in a multicast environment, such as a video server providing video (e.g., movies) to users via a computer network such as the Internet. The users may receive the data via an appliance such as a personal computer, Internet-capable set-top box, or the like. In this environment, multiple clients require different types of service due to variations in their processing power, memory resources and available bandwidth. The server would like to provide different resolutions and/or qualities of the same video sequence to satisfy each type of client. Note that the server should provide content for all the service rates continuously, otherwise, an entire client type does not receive service.

This scenario is different from the unicast environment, where different levels of service are provided to account for dynamic changes in the point-to-point transmission. In this case, the server can adaptively switch between bitstreams to provide service commensurate with the available resources. Therefore, only one bitstream is transmitted at any instance and it can be tailored to the target bit rate. This scheme provides high quality video, and benefits from low

decoder complexity since single-layer decoding is always used.

Different services may involve improving the spatial resolution, temporal resolution and/or quality of the video transmission with increasing bit rates. This discussion focuses on providing different levels of spatial resolution. Consider the following example with two levels of service. Clients at the lower service rate receive QCIF resolution video and those at the higher service rate receive CIF resolution video. Both service rates receive video with the same temporal resolution.

One approach to providing multicast service is to simulcast code the sequence at all of the service rates. This approach produces high quality video at all service rates. However, since service must be provided to all client types, the encoder must transmit a large amount of data, i.e., the sum of all the service rates. The main expense in multicast transmission is the total used bandwidth (in terms of total number of packets transmitted). Network congestion is also a problem, so it may be necessary to constrain the total combined rate of all the service rates. The minimum bandwidth necessary is the largest single service rate that is able to provide adequate service to all clients. The following discussion assumes that a constraint on the total bandwidth is imposed where the constrained rate is between the

largest single service rate and the sum of all the service rates.

The simulcast approach can still be used to provide two levels of service by reducing the higher bitstream to the difference between the constrained total and the lower bitstream. Another approach is to use spatially scalability. Intuitively, one expects spatial scalability to perform better since the information in the base layer (QCIF) sequence is used to assist construction of the enhancement layer (CIF) sequence. In general, this is true, but it has been found in connection with the present invention that simulcast coding can outperform spatial scalability when a small proportion of bits are allocated to the base layer/lower-resolution stream. This is counterintuitive, since spatial scalability "reuses" information in the base layer and its enhancement layer has the same bit rate as the single layer bitstream used by simulcast coding. One explanation may be that the overhead incurred by using a scalable coding syntax surpasses the gain obtained. The base layer also may not provide good prediction when a small amount of bits is allocated to it.

FIGs 28 and 29 are examples of multicast broadcasting using simulcast coding and spatial scalability, respectively, when the total used bandwidth is constrained to 200 kbps and the user service rates are 50 kbps and 200 kbps.

In FIG. 28, a server 2800 provides simulcast-coded lower-resolution and higher-resolution streams at 40 and 160 kbps, respectively, to a 200 kbps channel 2810 and a switching device, such as a router 2820. The 40 kbps stream is routed to the first service via a path 2830, and the 160 kbps stream is routed to the second service via a path 2850. The router 2820 does not route the 160 kbps stream to the first service since the first service cannot handle this data rate. Moreover, the router 2820 does not route the 40 kbps stream to the second service since this service only has use for one of the streams, and can handle the higher resolution 160 kbps stream.

Any known computer-network routing protocol may be used to achieve this result. In particular, the router 2820 should be informed of which services can handle which data rates. Each service can represent many end users.

In FIG. 29, a server 2900 provides spatial scalability-coded base and enhancement layers at 40 and 160 kbps, respectively, to a 200 kbps channel 2910 and a switching device, such as a router 2920. The 40 kbps base layer is routed to the first service via a path 2930, and both the 40 kbps base layer and 160 kbps enhancement layer are routed to the second service via paths 2940 and 2950, respectively (which may be the same path). The second service receives both the

layers since they must be used together to obtain the maximum information (e.g., image resolution).

Note that the video for the 50 kbps service rate is identical regardless of which coding approach is used. That is, the user with the 50 kbps service receives only the 40 kbps lower-resolution stream when simulcast coding is used, or the equivalent 40 kbps base layer when scalability coding is used.

The 200 kbps service in the simulcast coding approach (FIG. 28) constructs video for the higher service rate using the 160 kbps stream. This service does not use the 40 kbps single layer stream since it is independent from the 160 kbps layer and therefore there is no benefit to using it.

However, the 200 kbps service in the spatial scalability approach (FIG. 29) can use the 40 kbps base layer in addition to the 160 kbps enhancement layer, allowing it to construct video for the higher service rate using 200 kbps. These figures support the intuition that spatial scalability can outperform simulcast coding.

As discussed, FIG. 22 shows the results of encoding the QCIF and CIF Carphone sequences using both spatial scalability and simulcast coding with different bit allocations, but a fixed total bandwidth of 0.32 Mbps. Note that spatial scalability outperforms simulcast coding for many different bit allocations. However, simulcast coding is more efficient if less

than about 20% of the total bandwidth is allocated to the base layer/lower-resolution stream.

FIG. 30 shows a decision boundary 2620 (from FIG. 26) between simulcast coding and spatial scalability for the Carphone QCIF and CIF sequences. In accordance with the invention, operating points above the decision boundary should use simulcast coding, and points below it should use spatial scalability coding. The definition of operating point used here is the average bitrate over the entire sequence. Note that similar analysis can be performed using a smaller subset of the sequence such as a group of pictures. The methods described in this document can be used to construct this curve and then used to determine whether simulcast coding or spatial scalability should be used to encode the video.

A fixed total bitrate gives only one constraint on two variables (the base and enhancement bit rates), therefore, one can use any distribution of the total bitrate. For example, assume that 0.5 Mbps total bitrate is available. In one case, 0.25 Mbps is used for the QCIF resolution, and 0.25 Mbps is used for the CIF resolution. Since this operating point 3000 is below the curve 2620, spatial scalability should be used to obtain the best quality for the CIF data. In another case, 0.1 Mbps is used for the QCIF resolution, and 0.4 Mbps is used for the CIF resolution. Since this operating point 3020 is above the curve 2620,

simulcast coding should be used to achieve the best quality for the CIF data. Recall that the quality of the QCIF data is the same regardless of whether simulcast coding or scalability coding is used.

While the simulations described here used a number of encoding and decoding experiments before settling on which transmission mode to use, one may be able to determine model parameters without having to run encoding/decoding experiments for every sequence.

Determining new model parameters for each sequence may not be necessary if one already has good model parameters. This may be the case for sequences that are similar. For example, in FIG. 26, the sequences carphone 2620, news 2640, and silentvoice 2650 have a similar coding decision boundary. Using the boundary obtained with one of these sequences to assist the simulcast/scalable coding decision may still be optimal for the coding of a different sequence, as long as the data points are not close to the decision boundary. Determining new model parameters for each sequence may be desirable, especially if one wants to maximize the quality of video delivery, despite the increased computational costs.

FIG. 31 illustrates an adaptive simulcast/spatial scalability encoder apparatus in accordance with the present invention. A higher-resolution video sequence (such as CIF), and a lower-resolution video sequence (such as QCIF) are provided to an analysis function

3110, and to respective switches 3130 and 3140. The analysis function 3110 analyzes the video sequences based on the above discussion to provide a select signal to the switches 3130, 3140 to route both of the sequences to either the simulcast encoder 400 or the scalable encoder 100. See Figures 1 and 4, respectively.

The analysis function 3110 may include a decoder and encoder, where the PSNR of a layer is determined by comparing the output of the decoder to the input to the encoder.

Successive sequences, each having several (e.g., 150) pictures/frames, may be analyzed to adaptively route each sequence to either the simulcast encoder 400 or the scalable encoder 100. Moreover, the analysis may occur off-line, prior to when the video is transmitted to a user. In this manner, unnecessary processing delays are avoided. For example, the video data may be recovered from a memory, analyzed, then returned to the storage device. Each sequence of analyzed data may be marked to indicate whether it is to be subsequently routed to either the simulcast encoder 400 or the scalable encoder 100. This marking may be achieved any number of ways, such as providing overhead control bits with the video data.

Moreover, note that the decoders should have the capability to determine whether they are receiving a simulcast-coded or scalability-coded stream. This can

be achieved according to the relevant data standard (e.g., MPEG-4). Moreover, the decoders should have scalable decoding capabilities and single layer decoding capabilities. Advantageously, the invention can be carried out without modification to such decoders.

It should now be appreciated that the present invention provides for the compression of video data for multicast environments. Spatial scalability and simulcast coding are used in the compression process. Simulations are provided to compare the performance of spatial scalability and simulcast coding of lower-resolution (e.g., QCIF) and higher-resolution (e.g., CIF) sequences. The main results and conclusions of this work are:

- The use of a logarithmic model to represent single layer coding results. This model can then be used to easily compute the bit allocation that achieves equal quality in both layers of simulcast coding.
- Allocating 40 ± 4 % (e.g., 36-44%) of the total bandwidth to the lower-resolution stream achieves equal quality in both streams of simulcast coding.
- Allocating less than 20% of the total bandwidth to the base layer for spatial scalable coding is inefficient.
- Allocating 45 ± 5 % (e.g., 40-50%) of the total bandwidth to the base layer achieves equal quality in both layers of spatial scalable coding.

- Spatial scalability may be the only option at relatively low enhancement bit rates, and simulcast coding may be the only option at relatively high enhancement bit rates, but spatial scalability usually gives higher quality in the common operating region.
- Decision boundaries can be generated to guide the decision between spatial scalability and simulcast coding.
- Simulcast coding can outperform spatial scalability when a small proportion of bits is allocated to the base layer/lower-resolution stream. This may be due to the overhead of the scalable coding syntax and/or bad prediction from the base layer.

Although the invention has been described in connection with various preferred embodiments, it should be appreciated that various modifications and adaptations may be made thereto without departing from the scope of the invention as set forth in the claims.

What is claimed is:

1. A method for coding video data, comprising the steps of:

modeling a quality of lower resolution video data, and of corresponding higher resolution video data, for varying bit allocations between the lower and higher resolution video data;

determining a crossover point of the modeled qualities wherein the lower and higher resolution video data have an approximately equal quality, and the corresponding bit allocation; and

coding the video data in accordance with the corresponding bit allocation.

2. The method of claim 1, wherein:

the video data coded in said coding step comprises a sequence of pictures.

3. The method of claim 1, wherein:

the corresponding bit allocation is an average over a sequence of pictures that comprise the lower and higher resolution video data.

4. The method of claim 1, wherein:

the video data coded in said coding step is different than the video data used in said modeling step.

5. The method of claim 1, wherein:

the video data coded in said coding step is the same as the video data used in said modeling step.

6. The method of claim 1, wherein:

the quality comprises a peak signal-to-noise ratio (PSNR).

7. The method of claim 1, wherein:

said modeling step uses a logarithmic model.

8. The method of claim 1, wherein:

the coded video data is coded in independent simulcast streams.

9. The method of claim 1, wherein:

the coded video data is coded using spatial scalability coding in a base layer and at least one enhancement layer.

10. The method of claim 1, wherein:

the crossover point is determined in accordance with a bisection algorithm.

11. The method of claim 1, wherein:

the crossover point is extrapolated.

12. A method for coding video data, comprising the steps of:

modeling a quality of lower resolution video data, and of corresponding higher resolution video data, for varying bit allocations between the lower and higher resolution video data, by using each of: (a) simulcast coding, and (b) spatial scalability coding; and

coding the video data at a desired bit allocation for the lower and higher resolution video data using one of the simulcast coding and the spatial scalability coding according to whichever results in a larger modeled quality at the desired bit allocation.

13. The method of claim 12, wherein:

the video data comprises a sequence of pictures.

14. The method of claim 12, wherein:

the desired bit allocation is an average over a sequence of pictures that comprise the lower and higher resolution video data.

15. The method of claim 12, wherein:

the video data coded in said coding step is different than the video data used in said modeling step.

16. The method of claim 12, wherein:

the video data coded in said coding step is the same as the video data used in said modeling step.

17. The method of claim 12, wherein:

the modeled qualities comprise a peak signal-to-noise ratio (PSNR).

18. The method of claim 12, wherein:

said modeling step uses a logarithmic model.

19. A method for coding video data, comprising the steps of:

(a) determining a bit allocation decision boundary for varying bit allocations for lower resolution video data;

wherein, for the varying bit allocations for the lower resolution data, the decision boundary designates a bit allocation for corresponding higher resolution video data;

(b) determining an operating point for the video data based on desired bit allocations for lower resolution video data thereof and corresponding higher resolution video data thereof;

(c) coding the video data using simulcast coding if the operating point indicates the desired bit allocation of the higher resolution video data of said

step (b) exceeds a bit allocation designated by the decision boundary; and

(d) coding the video data using simulcast coding if the operating point indicates the desired bit allocation of the higher resolution video data of said step (b) is less than a bit allocation designated by the decision boundary.

20. The method of claim 19, wherein:

the decision boundary corresponds to a bit allocation of approximately 20% of an available bandwidth to the lower resolution video data in said step (a) and approximately 80% of the available bandwidth to the higher resolution video data in said step (a).

21. The method of claim 19, wherein:

the video data comprises a sequence of pictures.

22. The method of claim 21, wherein:

the desired bit allocations in said step (b) are averages over the sequence of pictures.

23. The method of claim 19, wherein:

the video data coded in said coding steps is different from the video data used in said step (a).

24. The method of claim 19, wherein:

the video data coded in said coding steps is the same as the video data used in said step (a).

25. A method for coding video data comprising lower resolution data and corresponding higher resolution data, comprising the steps of:

(a) allocating a first portion of an available bandwidth to the lower resolution data, and a remainder of the available bandwidth to the higher resolution data; and

(b) coding the video data using simulcast coding if the first portion is less than approximately 20% of the available bandwidth, and using spatial scalability coding if the first portion is greater than approximately 20% of the available bandwidth.

26. A method for coding video data comprising lower resolution data and corresponding higher resolution data, comprising the step of:

allocating approximately 36-44% of an available bandwidth to the lower resolution data, and a remainder of the available bandwidth to the higher resolution data to achieve an approximately equal quality for the lower and higher resolution data when they are simulcast coded.

27. The method of claim 26, wherein:

the quality comprises a peak signal-to-noise ratio (PSNR).

28. A method for coding video data comprising lower resolution data and corresponding higher resolution data, comprising the step of:

allocating approximately 40-50% of an available bandwidth to the lower resolution data, and a remainder of the available bandwidth to the higher resolution data to achieve an approximately equal quality for the lower and higher resolution data when they are spatial scalability coded.

29. The method of claim 28, wherein:

the quality comprises a peak signal-to-noise ratio (PSNR).

30. A method for communicating video data comprising lower resolution data and corresponding higher resolution data to at least a first user via a network that adaptively uses spatial scalability coding and simulcast coding, comprising the step of:

configuring a switching device in the network to selectively forward at least one of the lower and higher resolution data to the first user according to which of the spatial scalability coding and the

simulcast coding is being used, and according to a data service capacity of the first user.

31. The method of claim 30, wherein:

when the simulcast coding is used, the switching device is configured to: (a) forward the higher resolution data to the first user if a data rate of the higher resolution data is within the data service capacity, and (b) to forward the lower resolution data to the first user if the data rate of the higher resolution data is not within the data service capacity, but a data rate of the lower resolution data is within the data service capacity.

32. The method of claim 30, wherein:

when spatial scalability coding is used, the switching device is configured to: (a) forward both the higher and lower resolution data to the first user if a combined data rate thereof is within the data service capacity, and (b) to forward the lower resolution data to the first user if the combined data rate is not within the data service capacity, but a data rate of the lower resolution data is within the data service capacity.

33. The method of claim 30, wherein:

the switching device comprises a router.

34. An apparatus for coding video data, comprising:

means for modeling a quality of lower resolution video data, and of corresponding higher resolution video data, for varying bit allocations between the lower and higher resolution video data;

means for determining a crossover point of the modeled qualities wherein the lower and higher resolution video data have an approximately equal quality, and the corresponding bit allocation; and

means for coding the video data in accordance with the corresponding bit allocation.

35. An apparatus for coding video data, comprising:

means for modeling a quality of lower resolution video data, and of corresponding higher resolution video data, for varying bit allocations between the lower and higher resolution video data, by using each of: (a) simulcast coding, and (b) spatial scalability coding; and

means for coding the video data at a desired bit allocation for the lower and higher resolution video data using one of the simulcast coding and the spatial scalability coding according to whichever results in a larger modeled quality at the desired bit allocation.

36. An apparatus for coding video data, comprising:

(a) means for determining a bit allocation decision boundary for varying bit allocations for lower resolution video data;

wherein, for the varying bit allocations for the lower resolution data, the decision boundary designates a bit allocation for corresponding higher resolution video data;

(b) means for determining an operating point for the video data based on desired bit allocations for lower resolution video data thereof and corresponding higher resolution video data thereof;

(c) means for coding the video data using simulcast coding if the operating point indicates the desired bit allocation of the higher resolution video data of said means (b) exceeds a bit allocation designated by the decision boundary; and

(d) means coding the video data using simulcast coding if the operating point indicates the desired bit allocation of the higher resolution video data of said means (b) is less than a bit allocation designated by the decision boundary.

37. An apparatus for coding video data comprising lower resolution data and corresponding higher resolution data, comprising:

(a) means for allocating a first portion of an available bandwidth to the lower resolution data, and a remainder of the available bandwidth to the higher resolution data; and

(b) means for coding the video data using simulcast coding if the first portion is less than approximately 20% of the available bandwidth, and using spatial scalability coding if the first portion is greater than approximately 20% of the available bandwidth.

38. An apparatus for coding video data comprising lower resolution data and corresponding higher resolution data, comprising:

means for allocating approximately 36-44% of an available bandwidth to the lower resolution data, and a remainder of the available bandwidth to the higher resolution data to achieve an approximately equal quality for the lower and higher resolution data when they are simulcast coded.

39. An apparatus for coding video data comprising lower resolution data and corresponding higher resolution data, comprising:

means for allocating approximately 40-50% of an available bandwidth to the lower resolution data, and a remainder of the available bandwidth to the higher resolution data to achieve an approximately equal

quality for the lower and higher resolution data when they are spatial scalability coded.

40. An apparatus for communicating video data comprising lower resolution data and corresponding higher resolution data to at least a first user via a network that adaptively uses spatial scalability coding and simulcast coding, comprising:

means for configuring a switching device in the network to selectively forward at least one of the lower and higher resolution data to the first user according to which of the spatial scalability coding and the simulcast coding is being used, and according to a data service capacity of the first user.

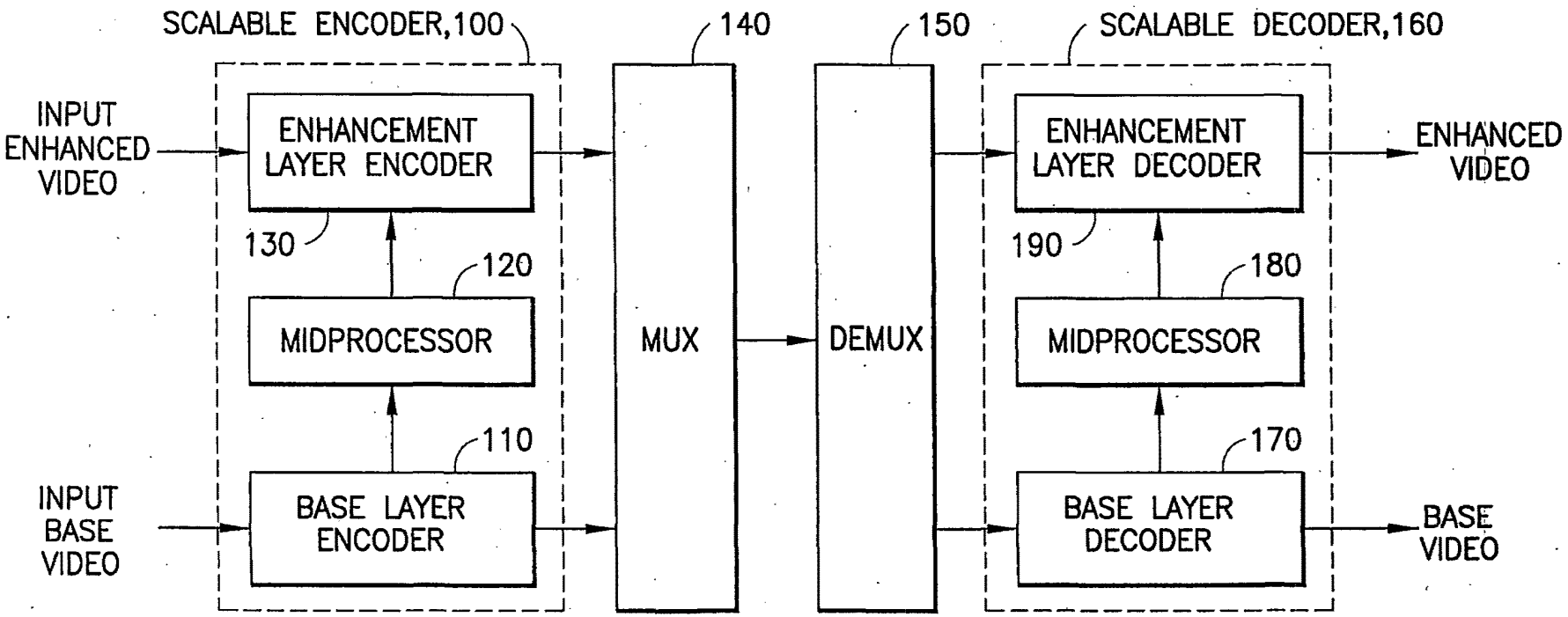

SCALABLE ENCODER,100
140
150
SCALABLE DECODER,160
INPUT ENHANCED VIDEO
ENHANCEMENT LAYER ENCODER
130
120
MIDPROCESSOR
110
INPUT BASE VIDEO
BASE LAYER ENCODER
MUX
DEMUX
ENHANCEMENT LAYER DECODER
190
180
MIDPROCESSOR
170
BASE LAYER DECODER
ENHANCED VIDEO
BASE VIDEO

FIG. 1

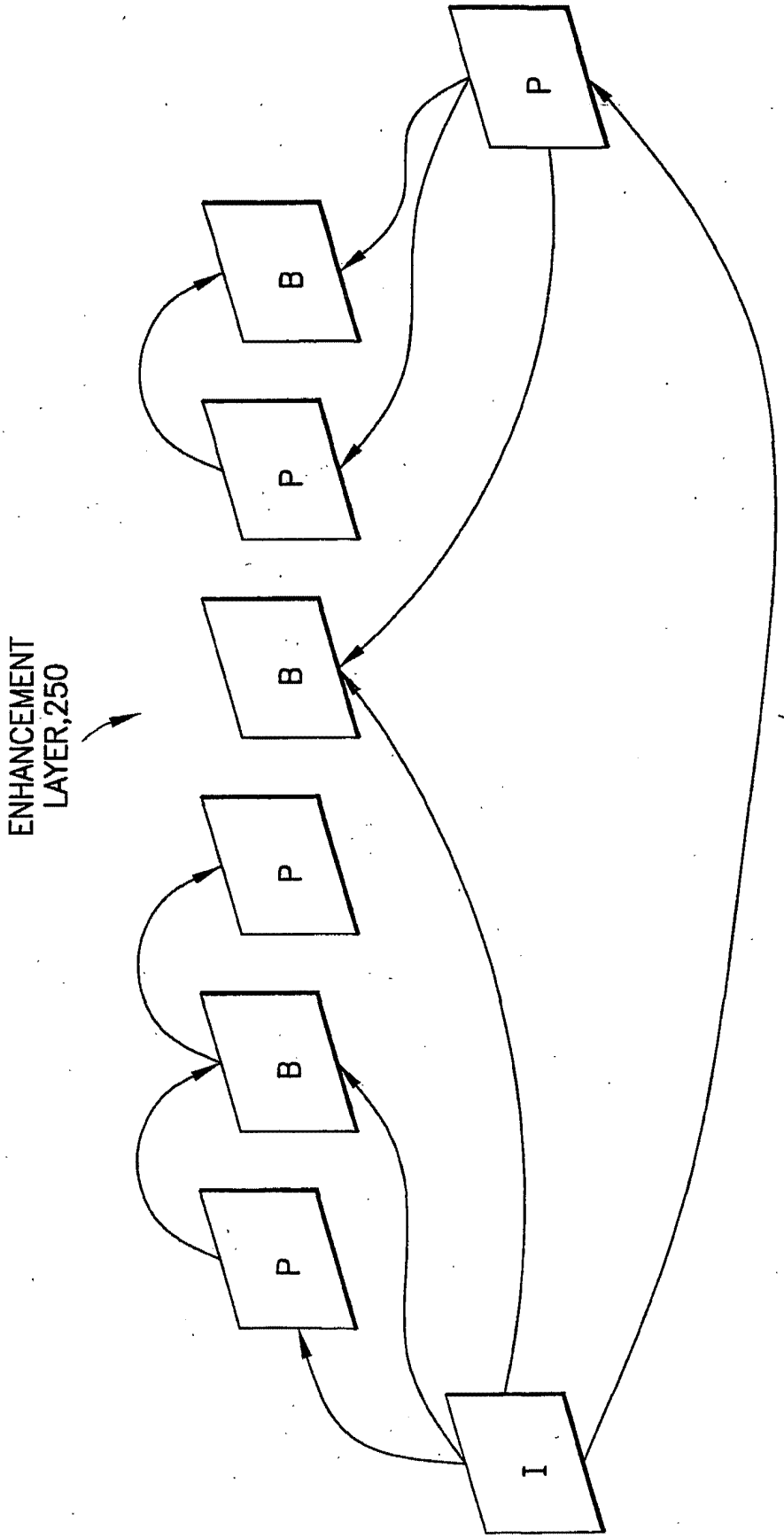


FIG.2

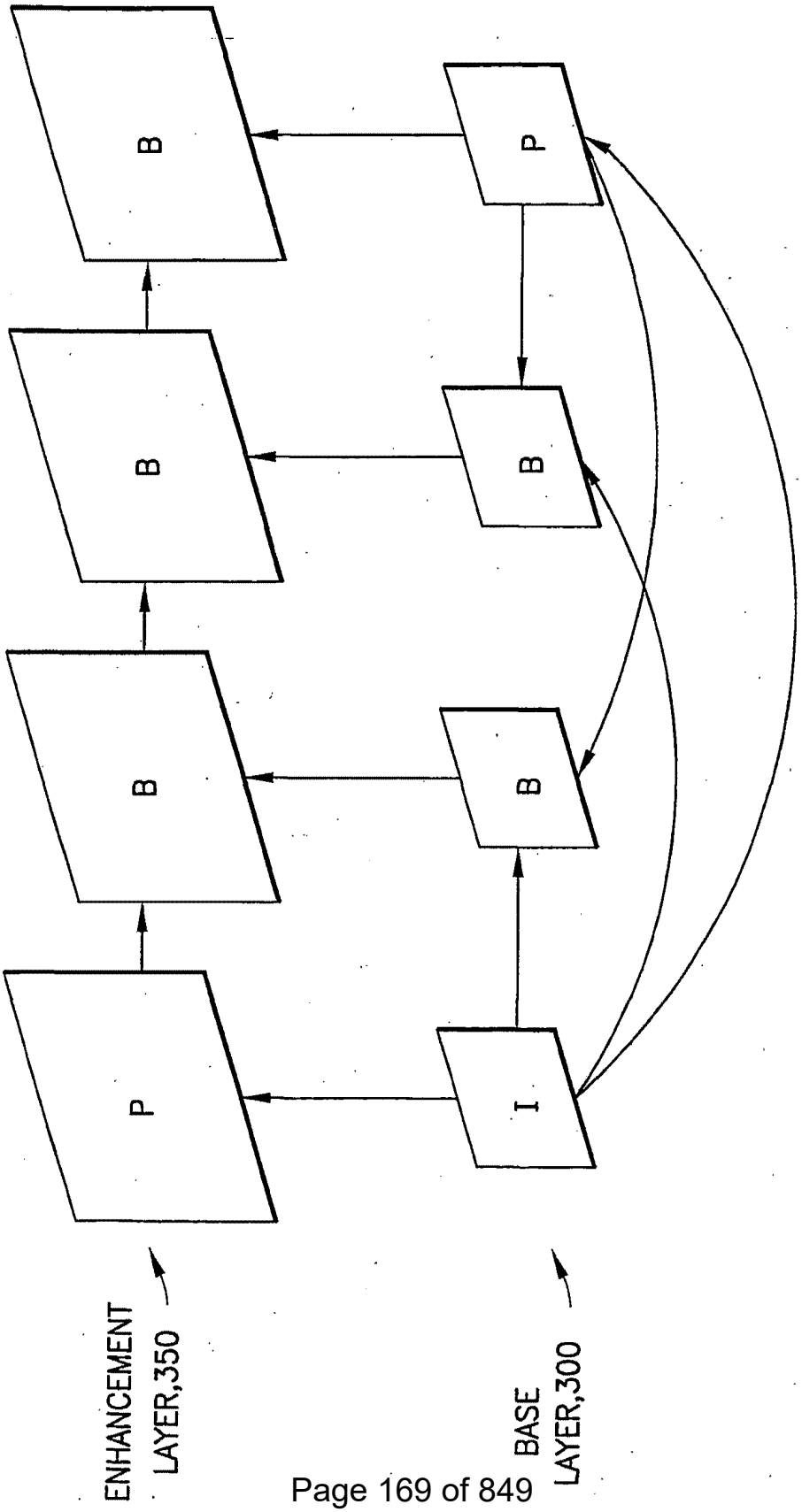


FIG.3

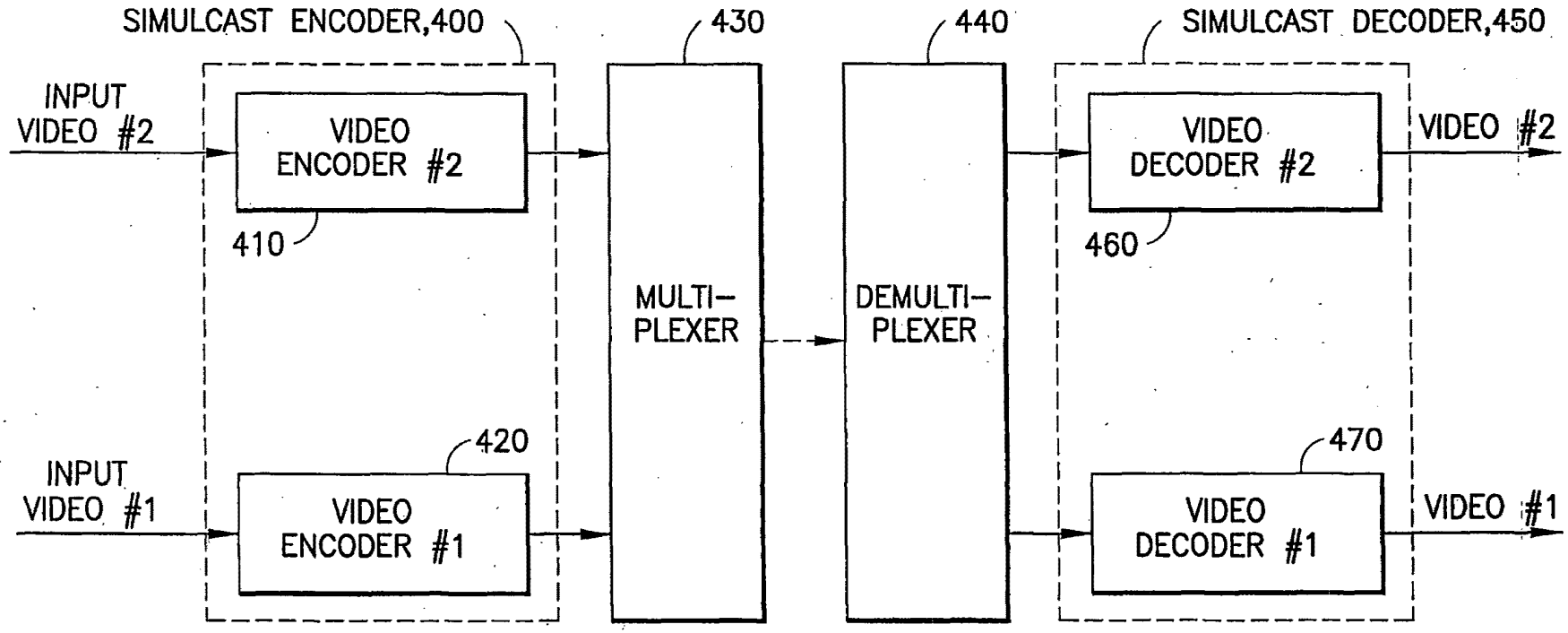

SIMULCAST ENCODER,400
430
440
SIMULCAST DECODER,450
INPUT VIDEO #2
VIDEO ENCODER #2
410
MULTI-PLEXER
DEMULTI-PLEXER
VIDEO DECODER #2
460
VIDEO #2
420
470
INPUT VIDEO #1
VIDEO ENCODER #1
VIDEO DECODER #1
VIDEO #1

FIG.4

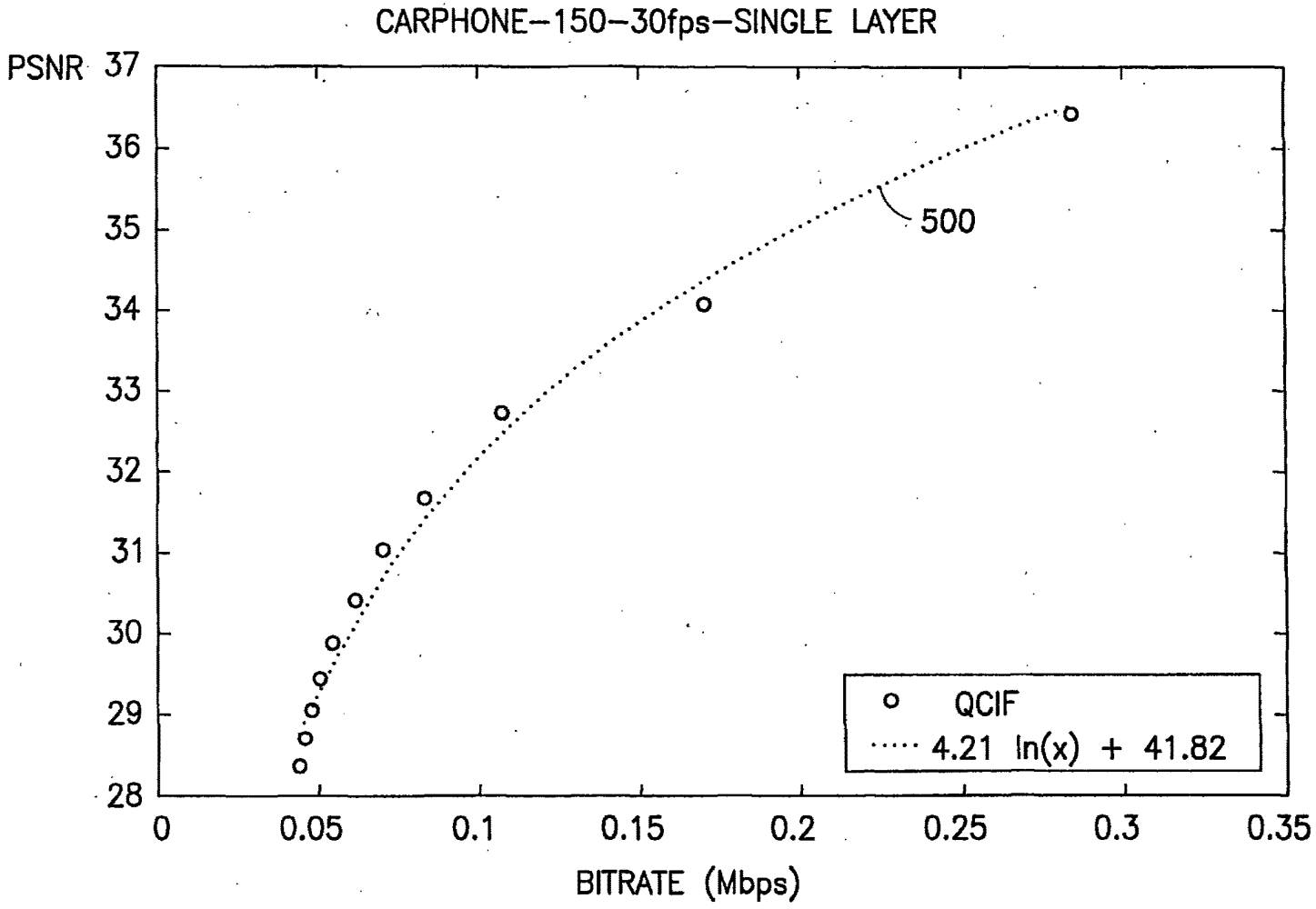

CARPHONE–150–30fps–SINGLE LAYER
PSNR
37
36
35
34
33
32
31
30
29
28
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
BITRATE (Mbps)
500
QCIF
4.21 ln(x) + 41.82

FIG.5

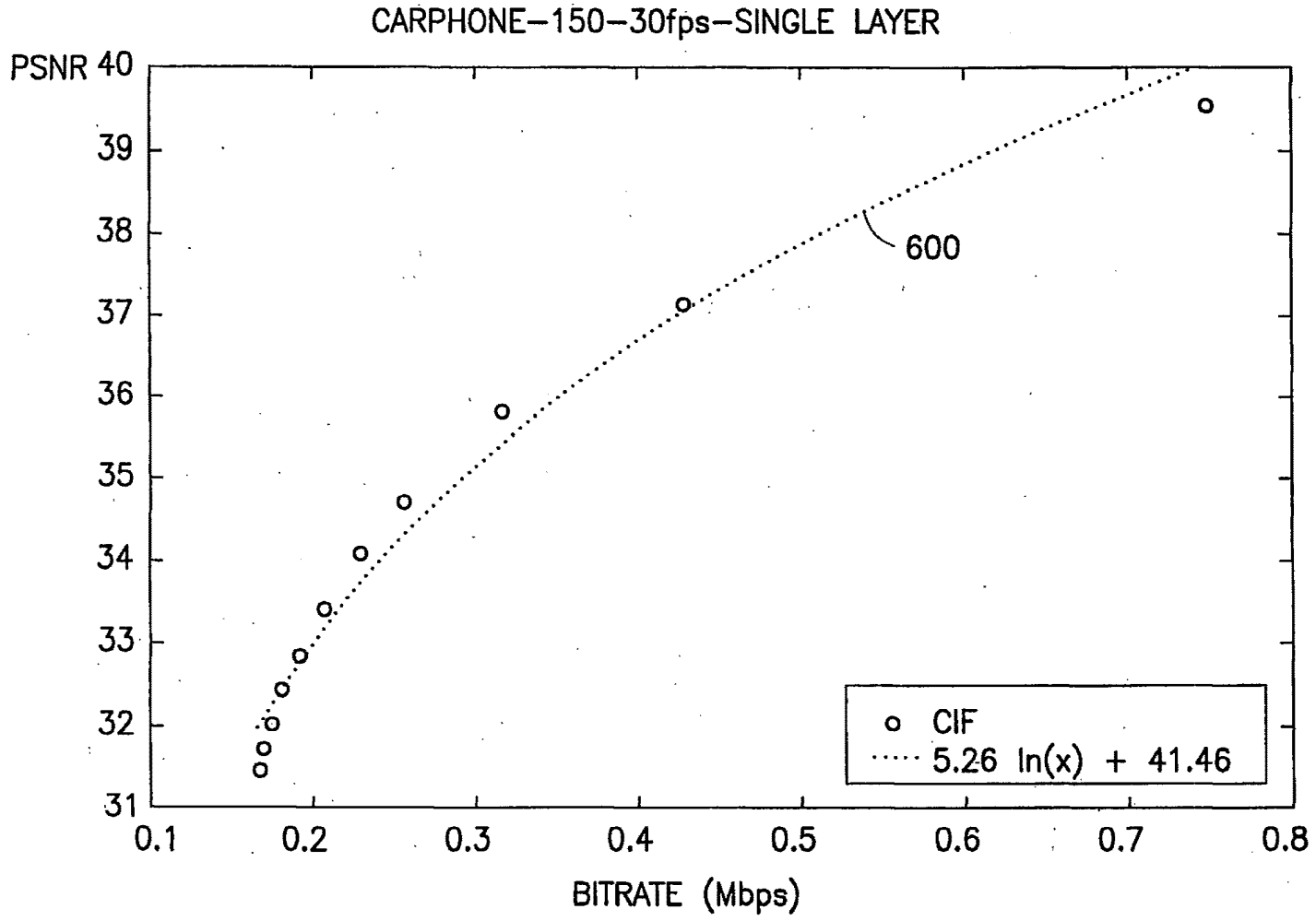

CARPHONE–150–30fps–SINGLE LAYER
PSNR
40
39
38
37
36
35
34
33
32
31
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
BITRATE (Mbps)
600
CIF
5.26 ln(x) + 41.46

FIG.6

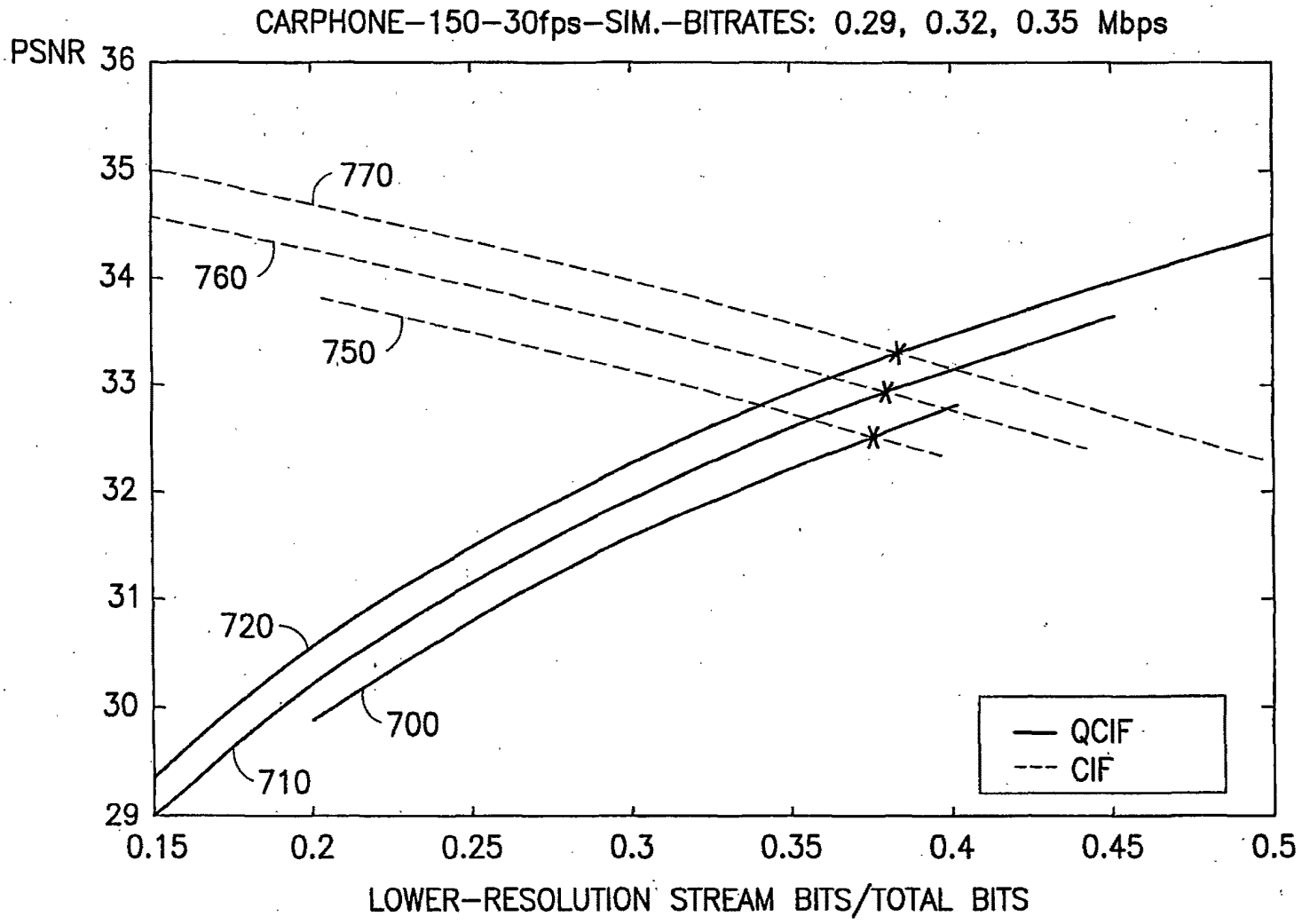

CARPHONE–150–30fps–SIM.–BITRATES: 0.29, 0.32, 0.35 Mbps
PSNR
36
35
34
33
32
31
30
29
770
760
750
720
700
710
QCIF
CIF
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5
LOWER–RESOLUTION STREAM BITS/TOTAL BITS

FIG.7

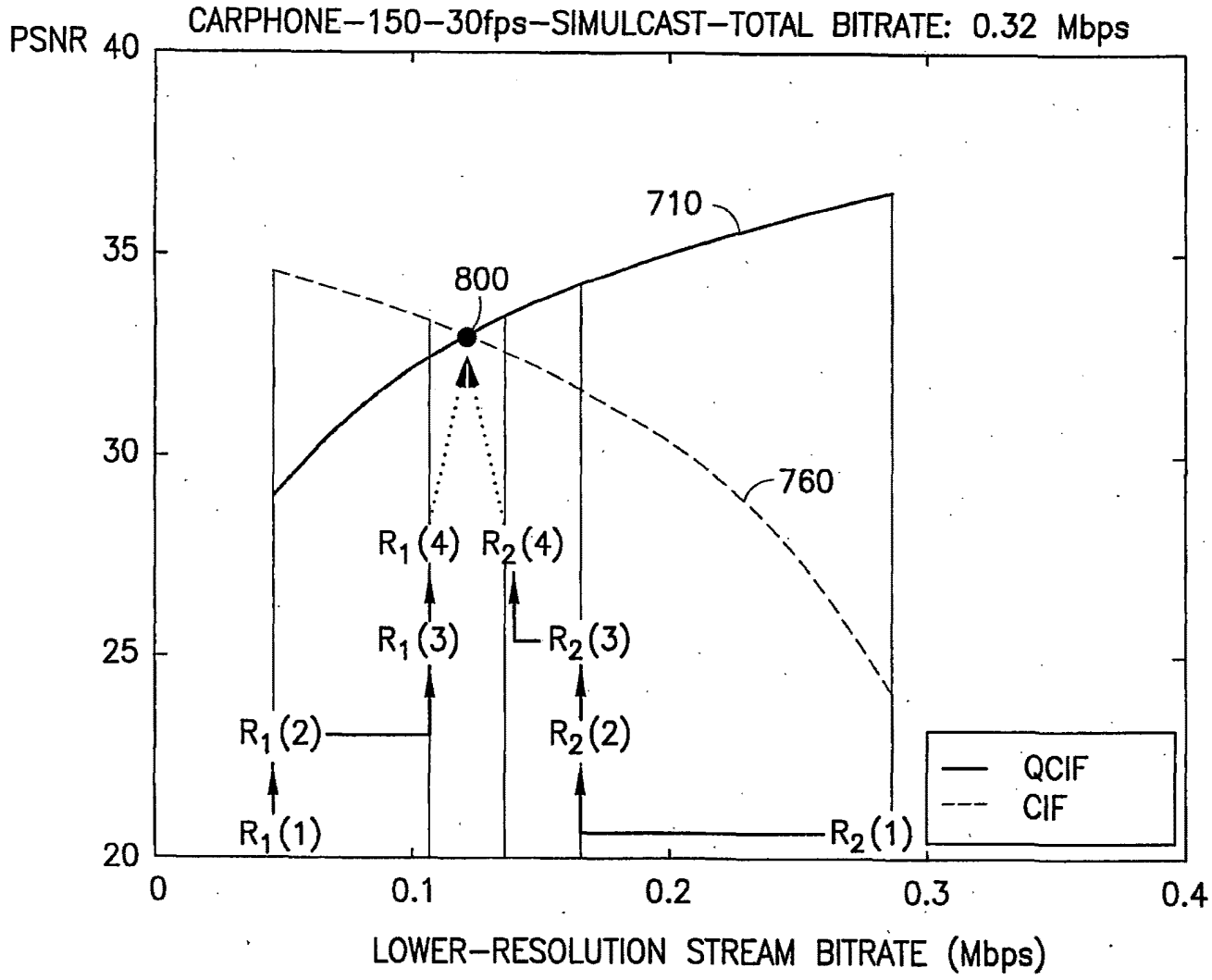

CARPHONE–150–30fps–SIMULCAST–TOTAL BITRATE: 0.32 Mbps
PSNR
40
35
30
25
20
0
0.1
0.2
0.3
0.4
LOWER–RESOLUTION STREAM BITRATE (Mbps)
710
800
760
R1(4)
R2(4)
R1(3)
R2(3)
R1(2)
R2(2)
R1(1)
R2(1)
QCIF
CIF

FIG.8

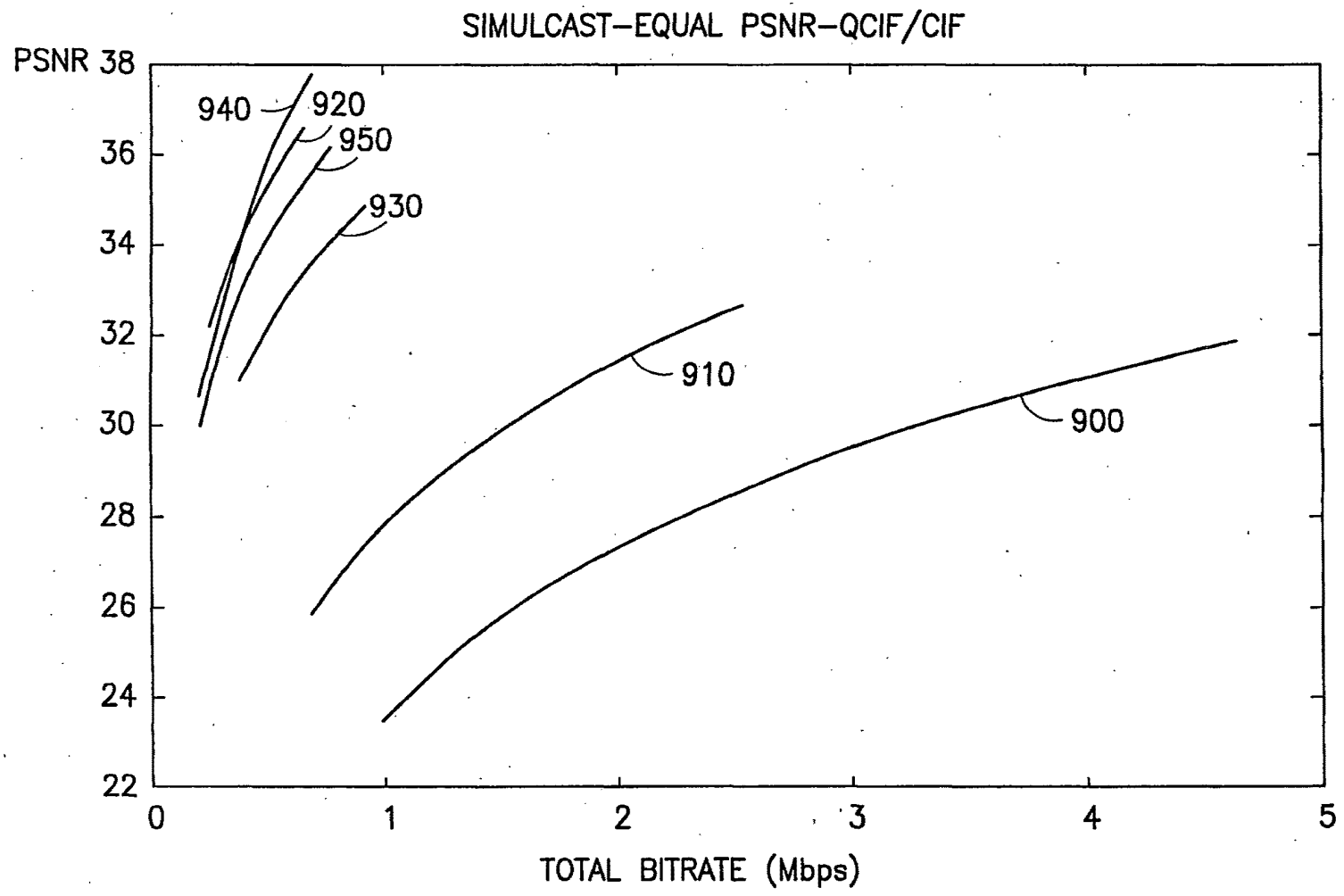

SIMULCAST–EQUAL PSNR–QCIF/CIF
PSNR
38
36
34
32
30
28
26
24
22
0
1
2
3
4
5
TOTAL BITRATE (Mbps)
940
920
950
930
910
900

FIG.9

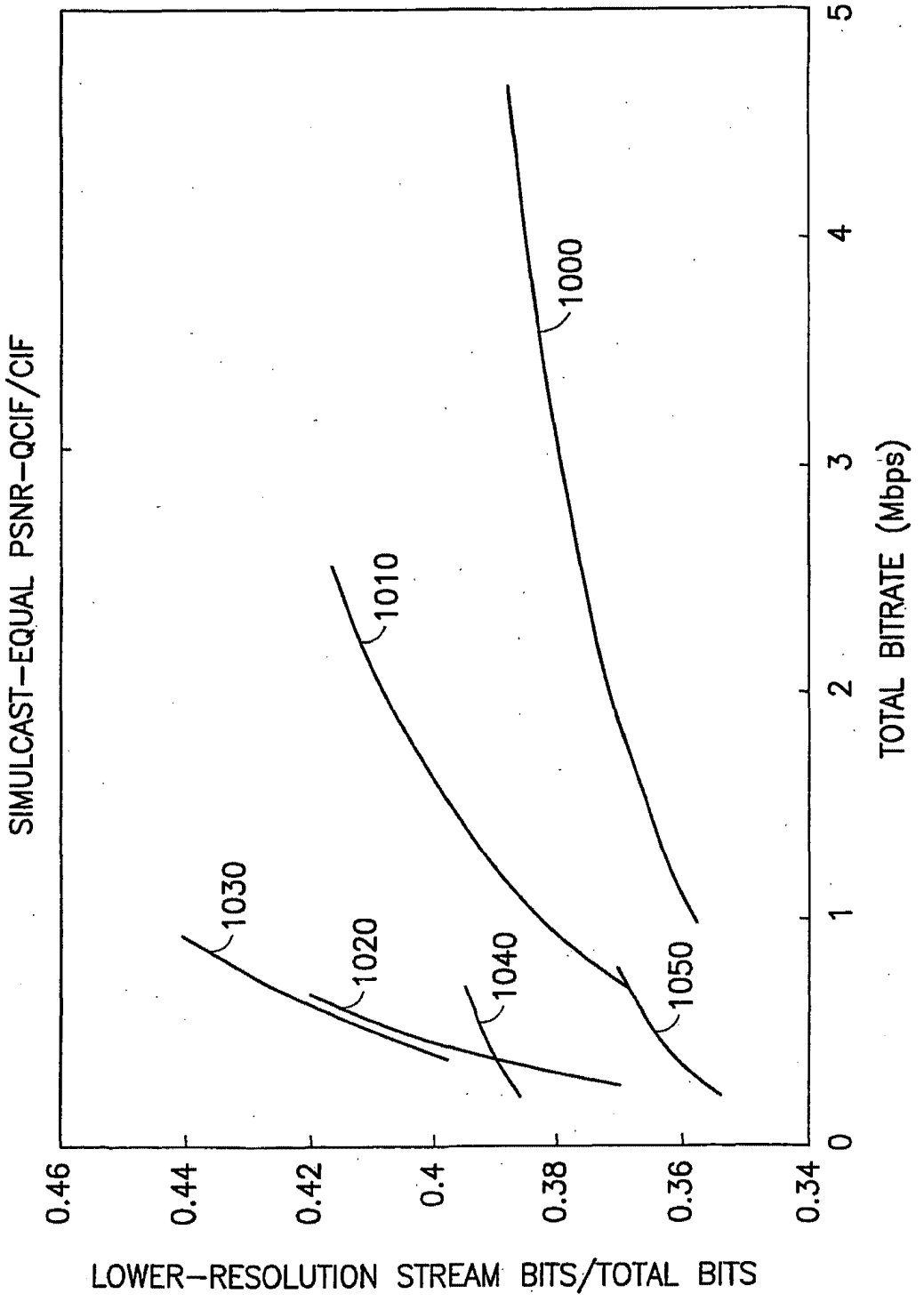

SIMULCAST–EQUAL PSNR–QCIF/CIF
LOWER–RESOLUTION STREAM BITS/TOTAL BITS
TOTAL BITRATE (Mbps)
0.46
0.44
0.42
0.4
0.38
0.36
0.34
0
1
2
3
4
5
1000
1010
1020
1030
1040
1050

FIG.10

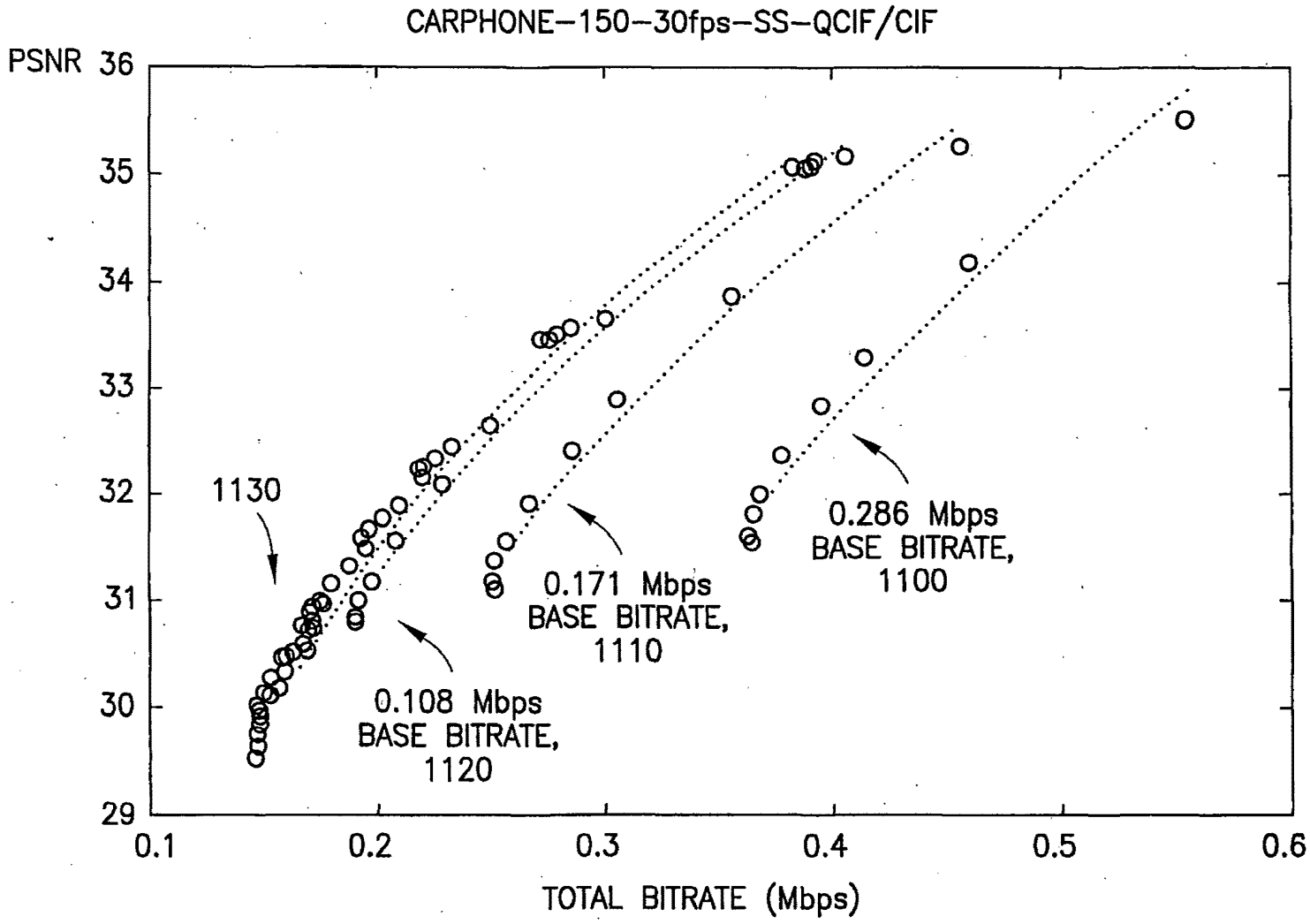

CARPHONE–150–30fps–SS–QCIF/CIF
PSNR
36
35
34
33
32
31
30
29
0.1
0.2
0.3
0.4
0.5
0.6
TOTAL BITRATE (Mbps)
1130
0.108 Mbps
BASE BITRATE,
1120
0.171 Mbps
BASE BITRATE,
1110
0.286 Mbps
BASE BITRATE,
1100

FIG.11

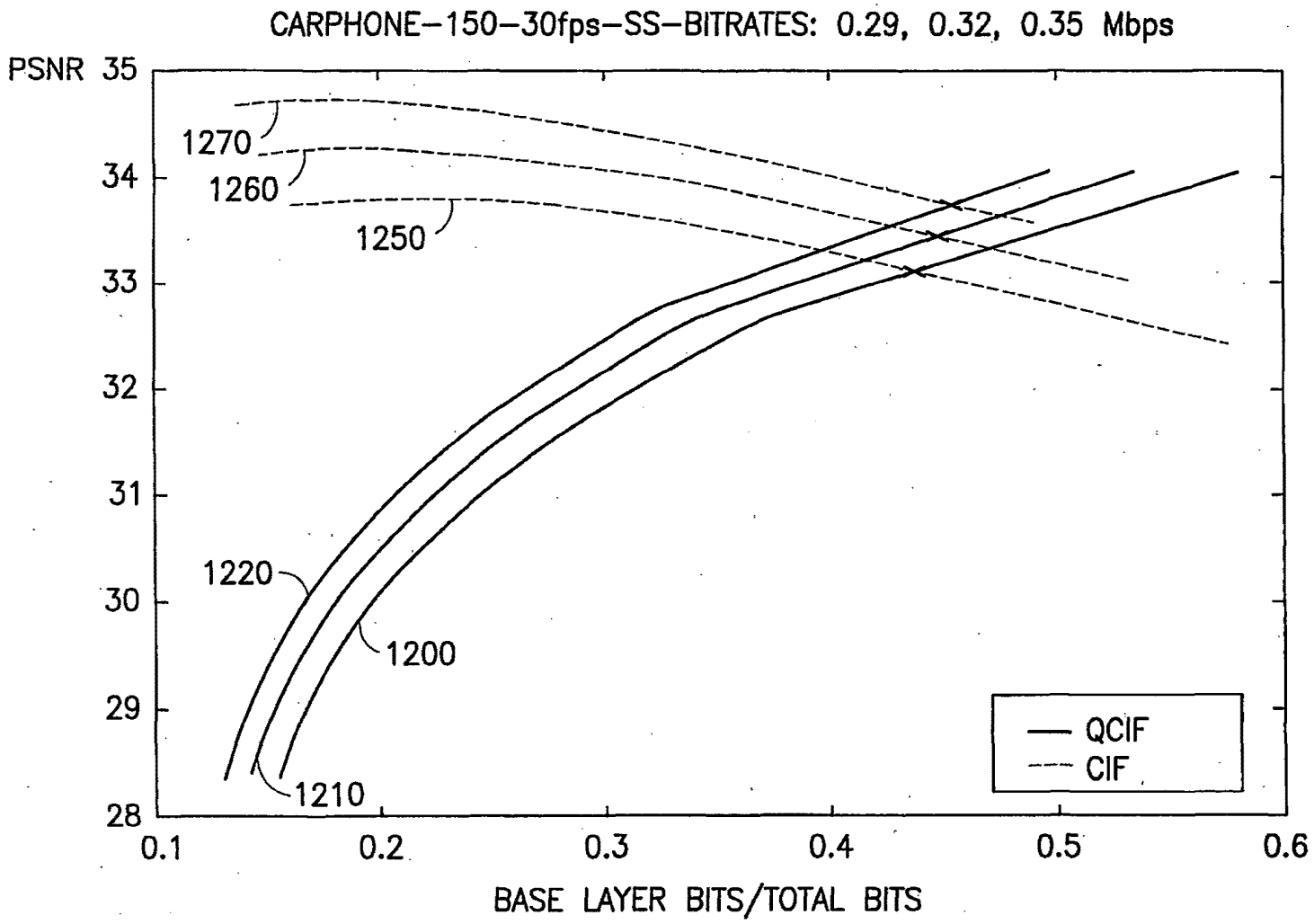

CARPHONE–150–30fps–SS–BITRATES: 0.29, 0.32, 0.35 Mbps
PSNR
35
34
33
32
31
30
29
28
0.1
0.2
0.3
0.4
0.5
0.6
BASE LAYER BITS/TOTAL BITS
1270
1260
1250
1220
1200
1210
QCIF
CIF

FIG.12

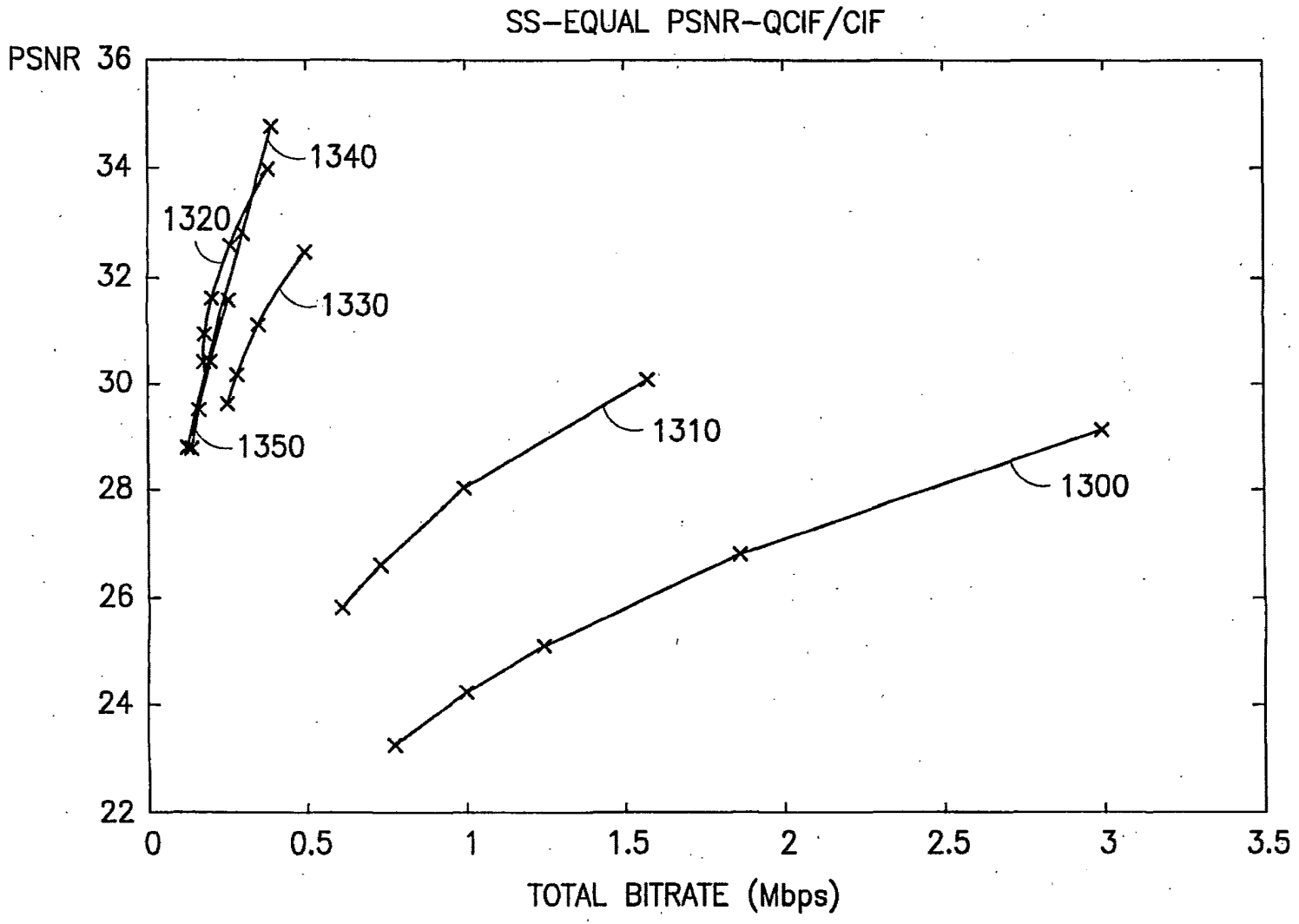

SS-EQUAL PSNR-QCIF/CIF
PSNR
36
34
32
30
28
26
24
22
0
0.5
1
1.5
2
2.5
3
3.5
TOTAL BITRATE (Mbps)
1340
1320
1330
1350
1310
1300

FIG.13

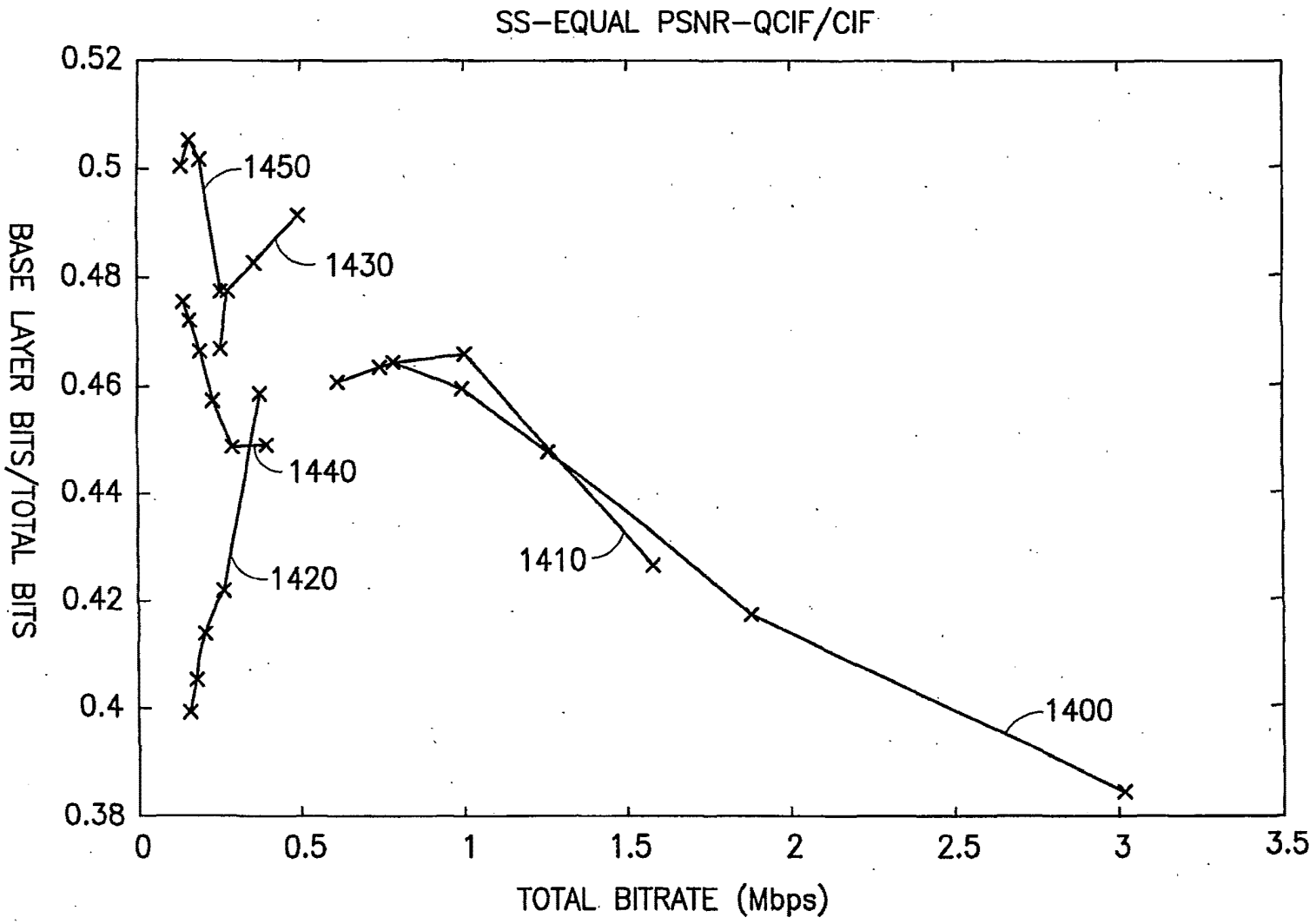

SS-EQUAL PSNR-QCIF/CIF
BASE LAYER BITS/TOTAL BITS
TOTAL BITRATE (Mbps)
0.52
0.5
0.48
0.46
0.44
0.42
0.4
0.38
0
0.5
1
1.5
2
2.5
3
3.5
1450
1430
1440
1420
1410
1400

FIG.14

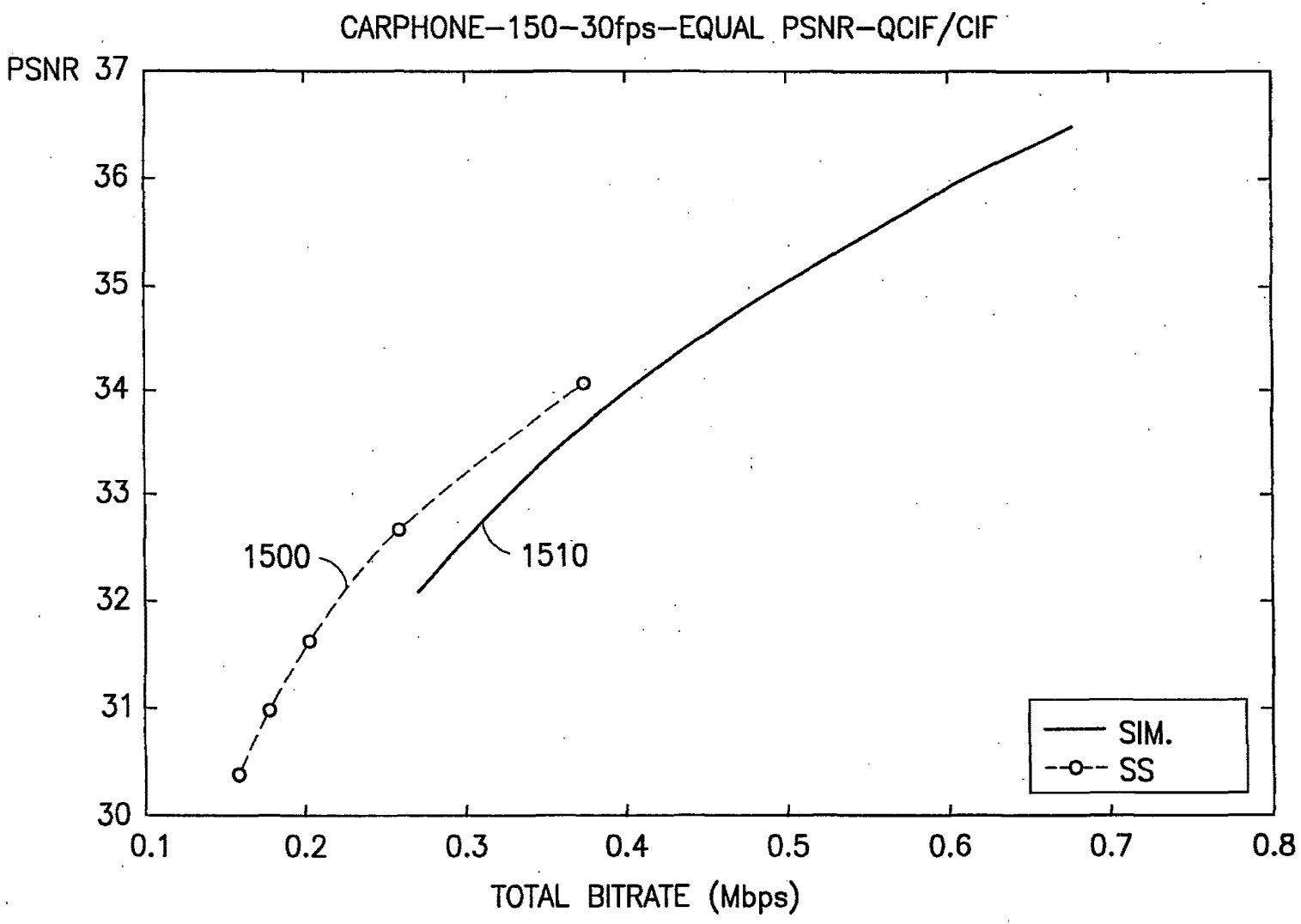

CARPHONE–150–30fps–EQUAL PSNR–QCIF/CIF
PSNR
37
36
35
34
33
32
31
30
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
TOTAL BITRATE (Mbps)
1500
1510
SIM.
SS

FIG.15

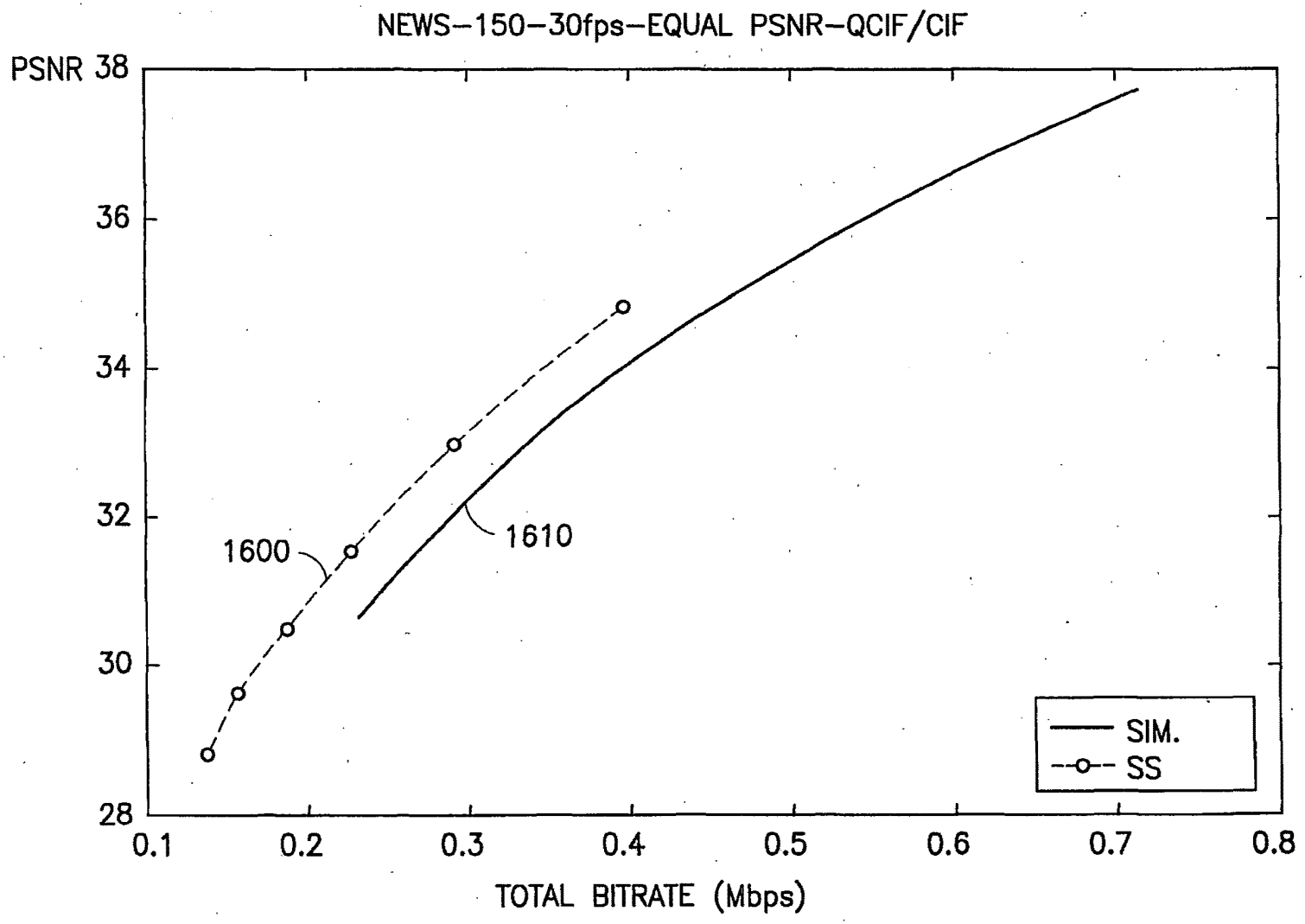

NEWS–150–30fps–EQUAL PSNR–QCIF/CIF
PSNR 38
36
34
32
30
28
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
TOTAL BITRATE (Mbps)
1600
1610
SIM.
SS

FIG.16

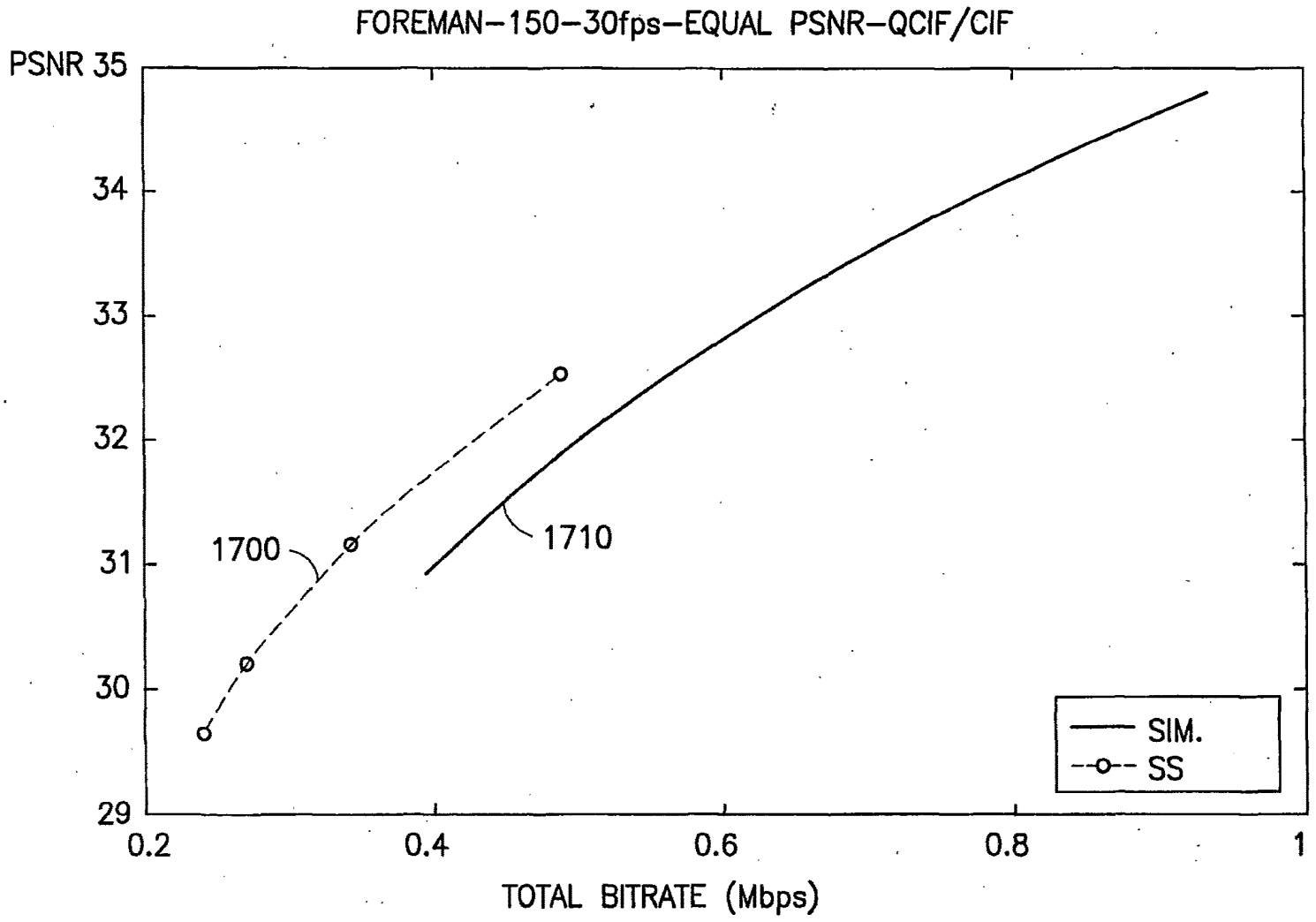

FOREMAN–150–30fps–EQUAL PSNR–QCIF/CIF
PSNR 35
34
33
32
31
30
29
0.2
0.4
0.6
0.8
1
TOTAL BITRATE (Mbps)
1700
1710
SIM.
SS

FIG.17

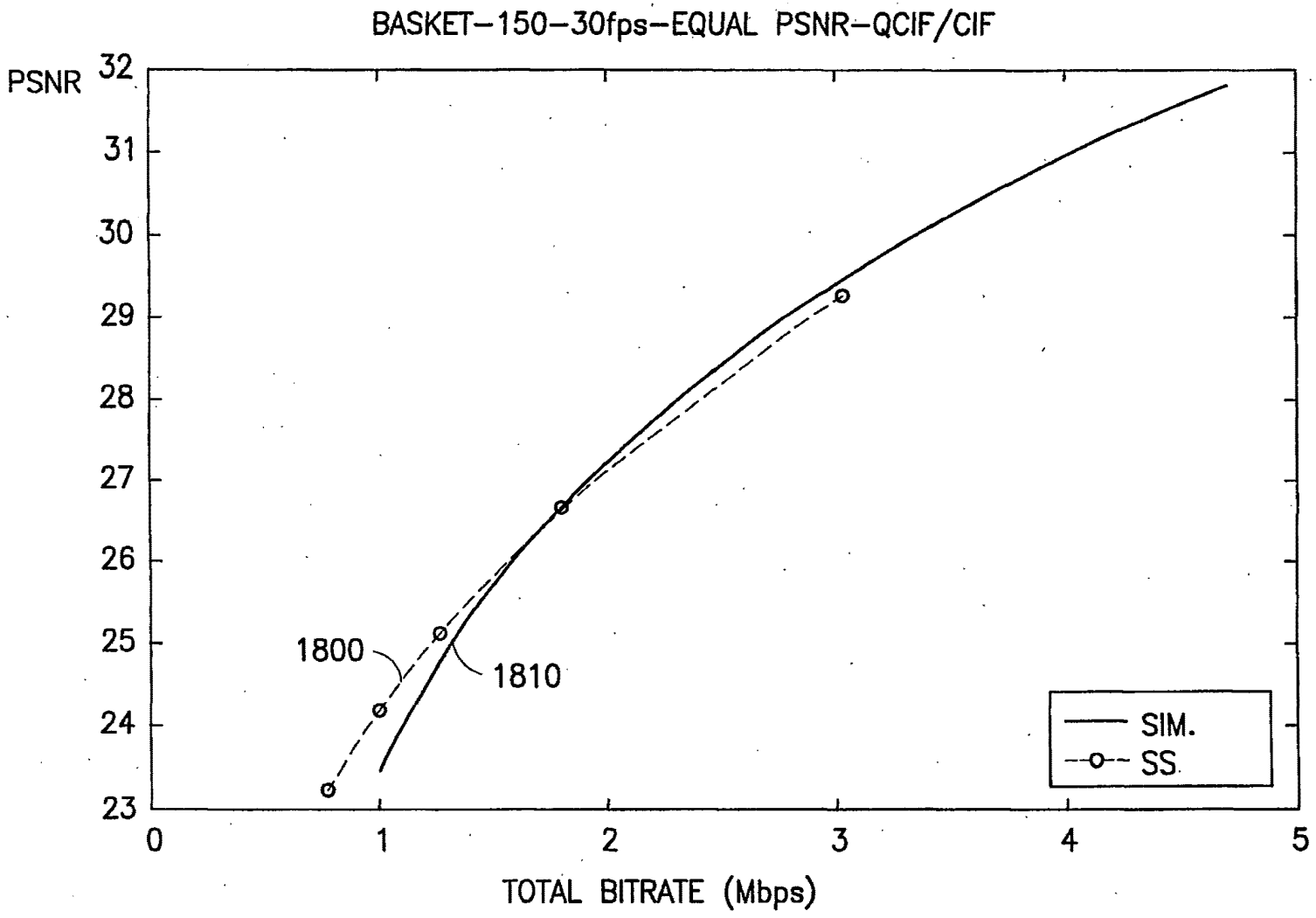

BASKET–150–30fps–EQUAL PSNR–QCIF/CIF
PSNR
32
31
30
29
28
27
26
25
24
23
0
1
2
3
4
5
TOTAL BITRATE (Mbps)
1800
1810
SIM.
SS

FIG.18

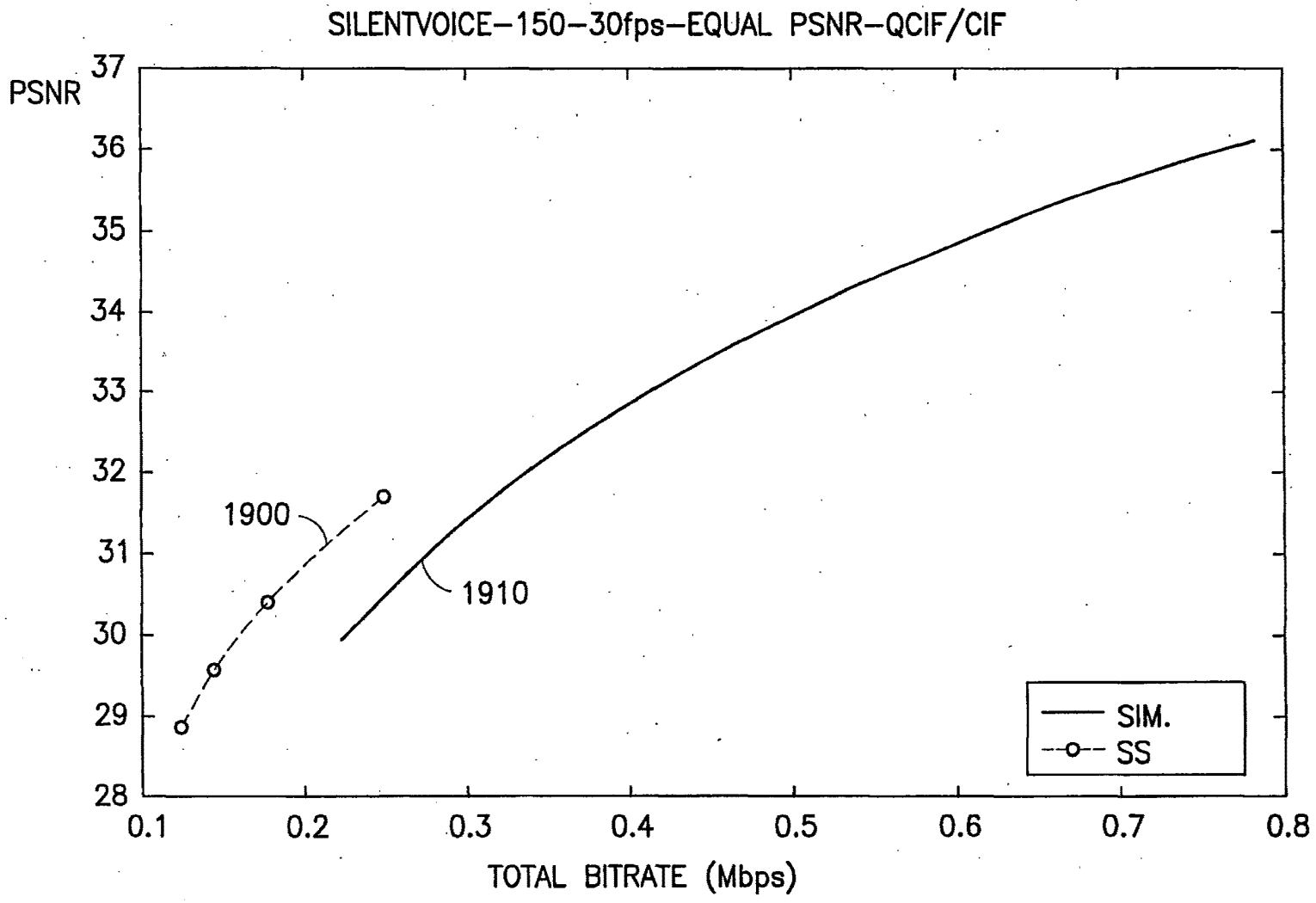

SILENTVOICE–150–30fps–EQUAL PSNR–QCIF/CIF
PSNR
37
36
35
34
33
32
31
30
29
28
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
TOTAL BITRATE (Mbps)
1900
1910
SIM.
SS

FIG.19

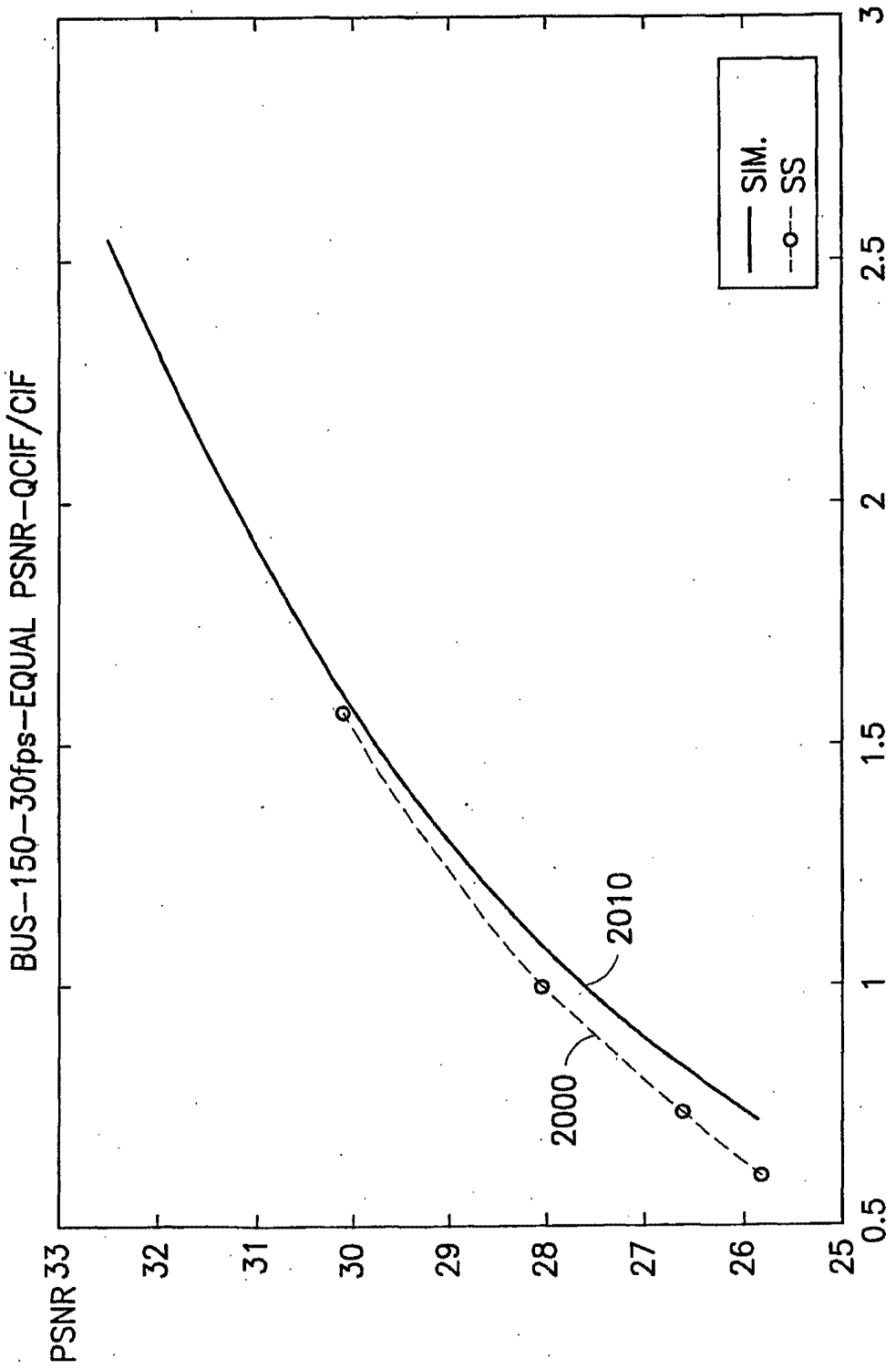

BUS-150-30fps-EQUAL PSNR-QCIF/CIF
PSNR
33
32
31
30
29
28
27
26
25
0.5
1
1.5
2
2.5
3
TOTAL BITRATE (Mbps)
SIM.
SS
2010
2000

FIG.20

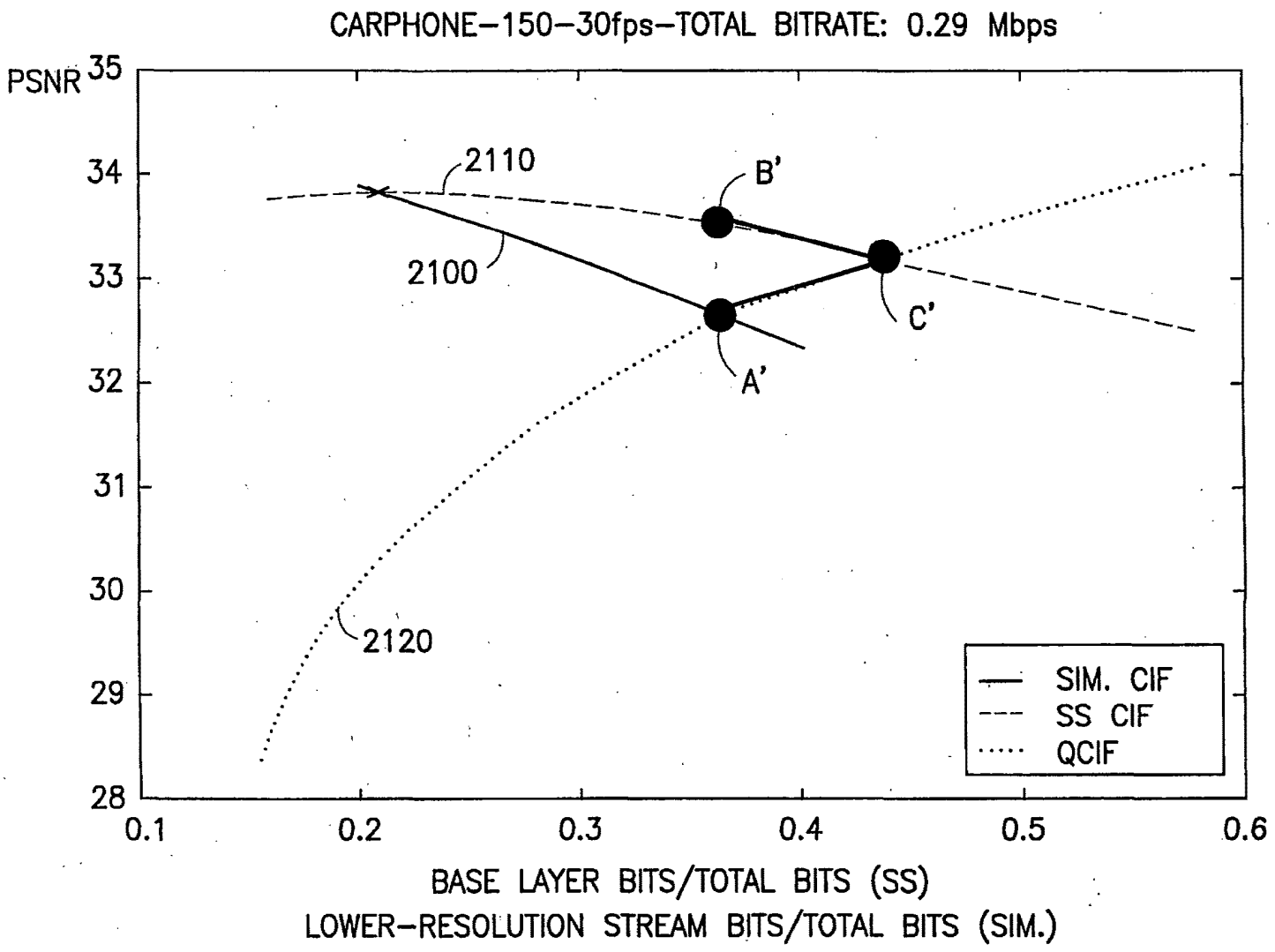

CARPHONE–150–30fps–TOTAL BITRATE: 0.29 Mbps
PSNR
35
34
33
32
31
30
29
28
0.1
0.2
0.3
0.4
0.5
0.6
2110
2100
2120
B'
C'
A'
SIM. CIF
SS CIF
QCIF
BASE LAYER BITS/TOTAL BITS (SS)
LOWER–RESOLUTION STREAM BITS/TOTAL BITS (SIM.)

FIG.21

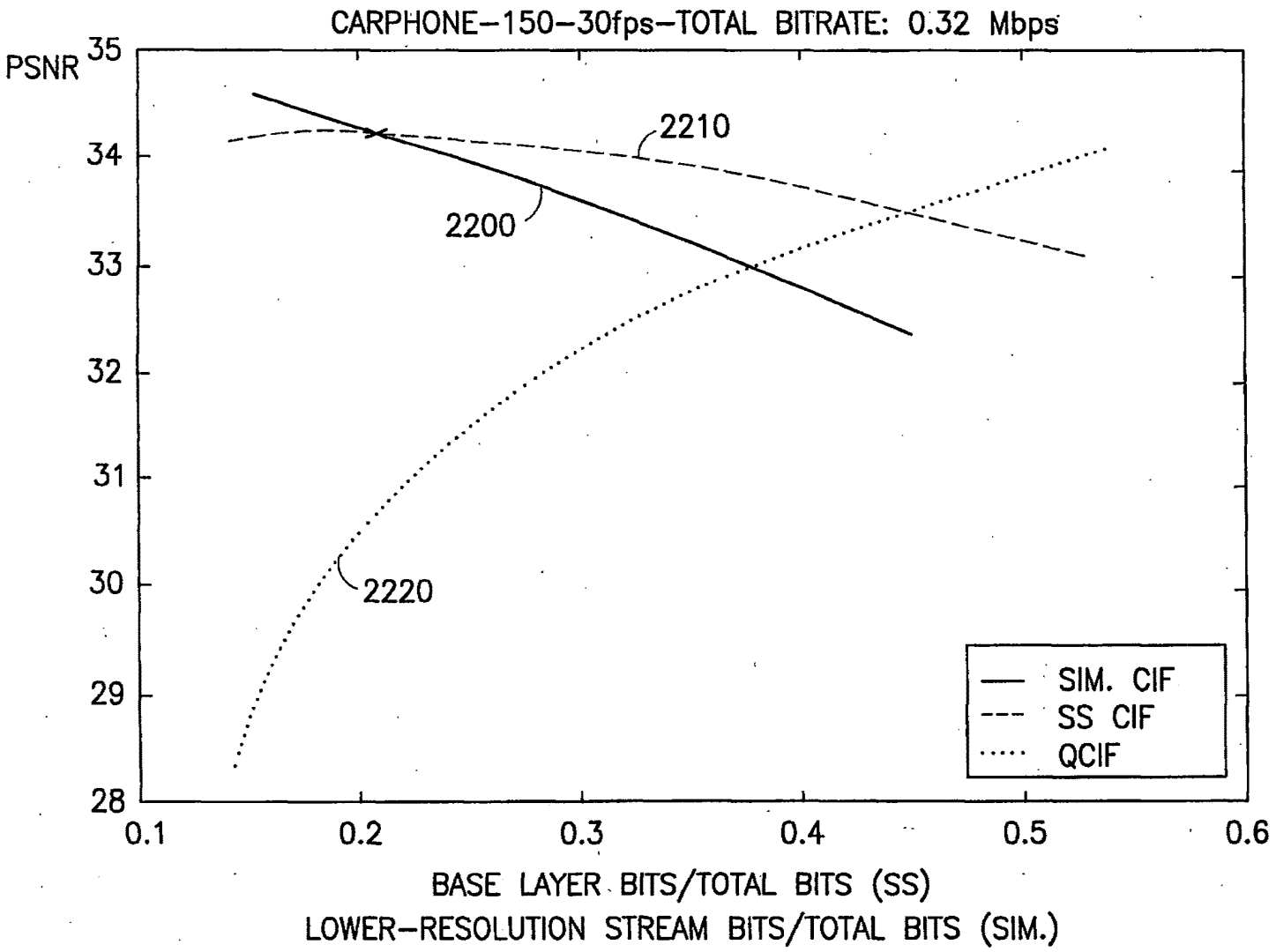

CARPHONE–150–30fps–TOTAL BITRATE: 0.32 Mbps
PSNR
35
34
33
32
31
30
29
28
0.1
0.2
0.3
0.4
0.5
0.6
2210
2200
2220
SIM. CIF
SS CIF
QCIF
BASE LAYER BITS/TOTAL BITS (SS)
LOWER–RESOLUTION STREAM BITS/TOTAL BITS (SIM.)

FIG.22

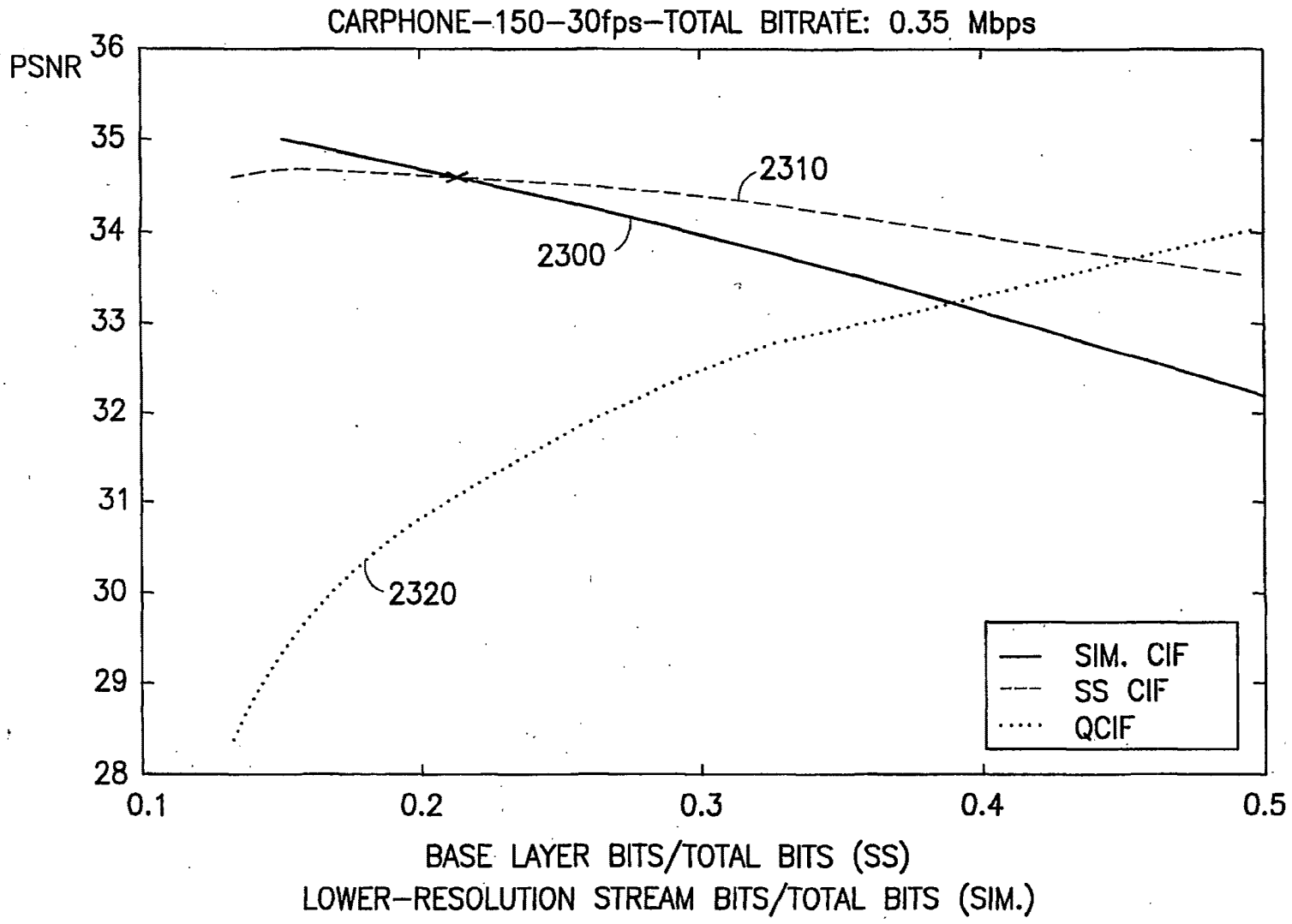

CARPHONE–150–30fps–TOTAL BITRATE: 0.35 Mbps
PSNR
36
35
34
33
32
31
30
29
28
0.1
0.2
0.3
0.4
0.5
2310
2300
2320
SIM. CIF
SS CIF
QCIF
BASE LAYER BITS/TOTAL BITS (SS)
LOWER–RESOLUTION STREAM BITS/TOTAL BITS (SIM.)

FIG.23

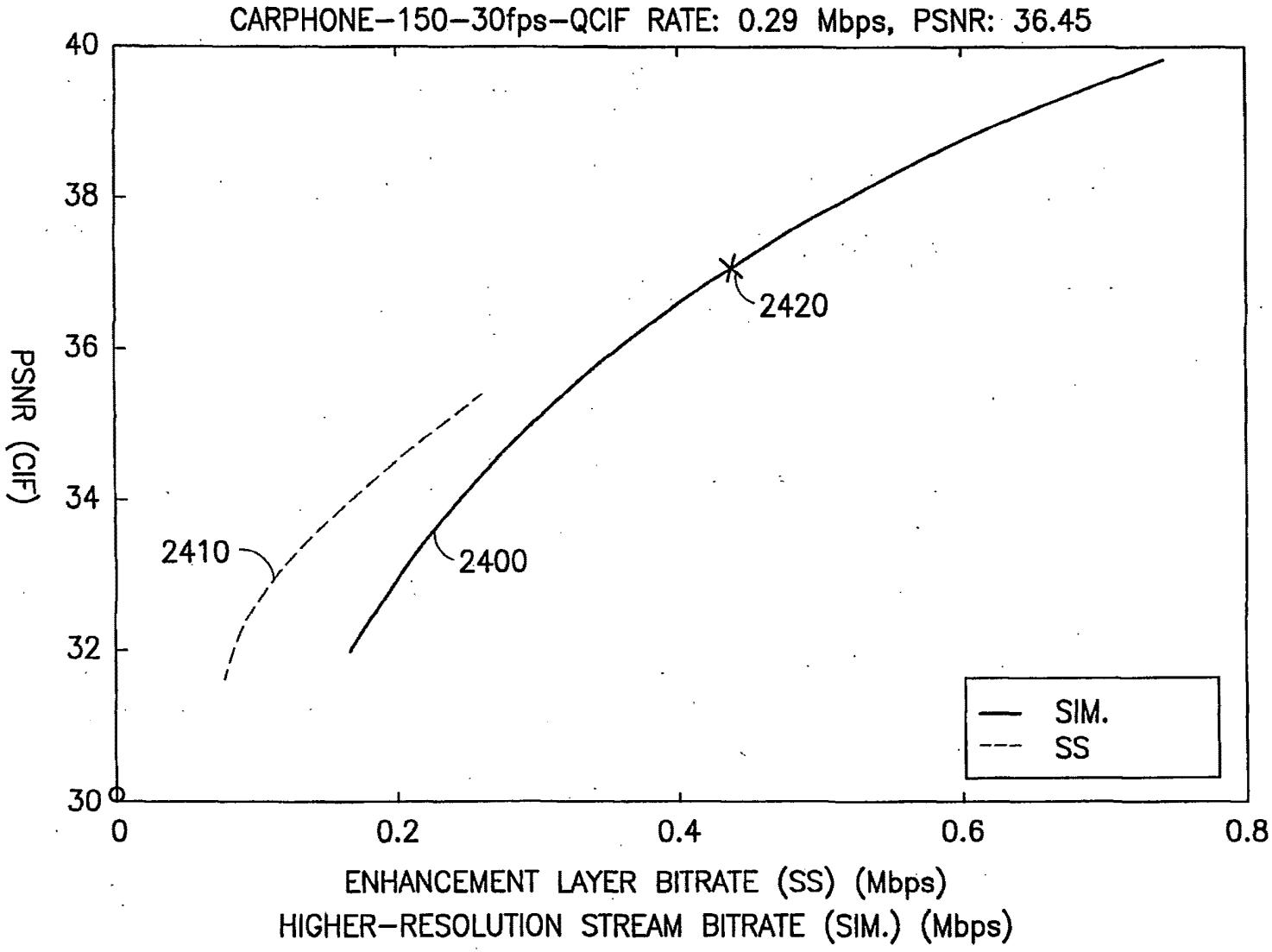

CARPHONE–150–30fps–QCIF RATE: 0.29 Mbps, PSNR: 36.45
PSNR (CIF)
40
38
36
34
32
30
0
0.2
0.4
0.6
0.8
2420
2410
2400
SIM.
SS
ENHANCEMENT LAYER BITRATE (SS) (Mbps)
HIGHER–RESOLUTION STREAM BITRATE (SIM.) (Mbps)

FIG.24

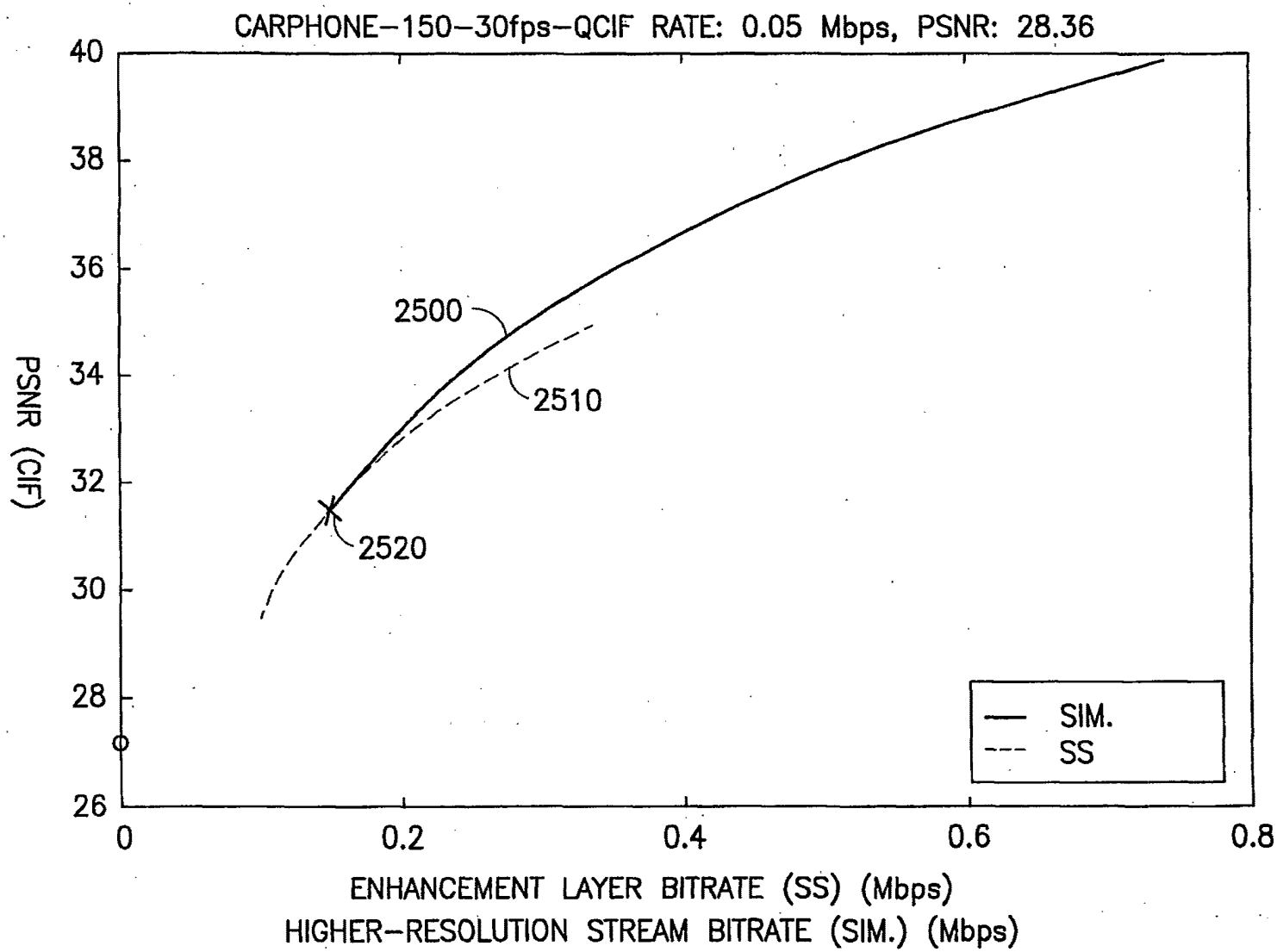

CARPHONE–150–30fps–QCIF RATE: 0.05 Mbps, PSNR: 28.36
PSNR (CIF)
40
38
36
34
32
30
28
26
0
0.2
0.4
0.6
0.8
2500
2510
2520
SIM.
SS
ENHANCEMENT LAYER BITRATE (SS) (Mbps)
HIGHER–RESOLUTION STREAM BITRATE (SIM.) (Mbps)

FIG.25

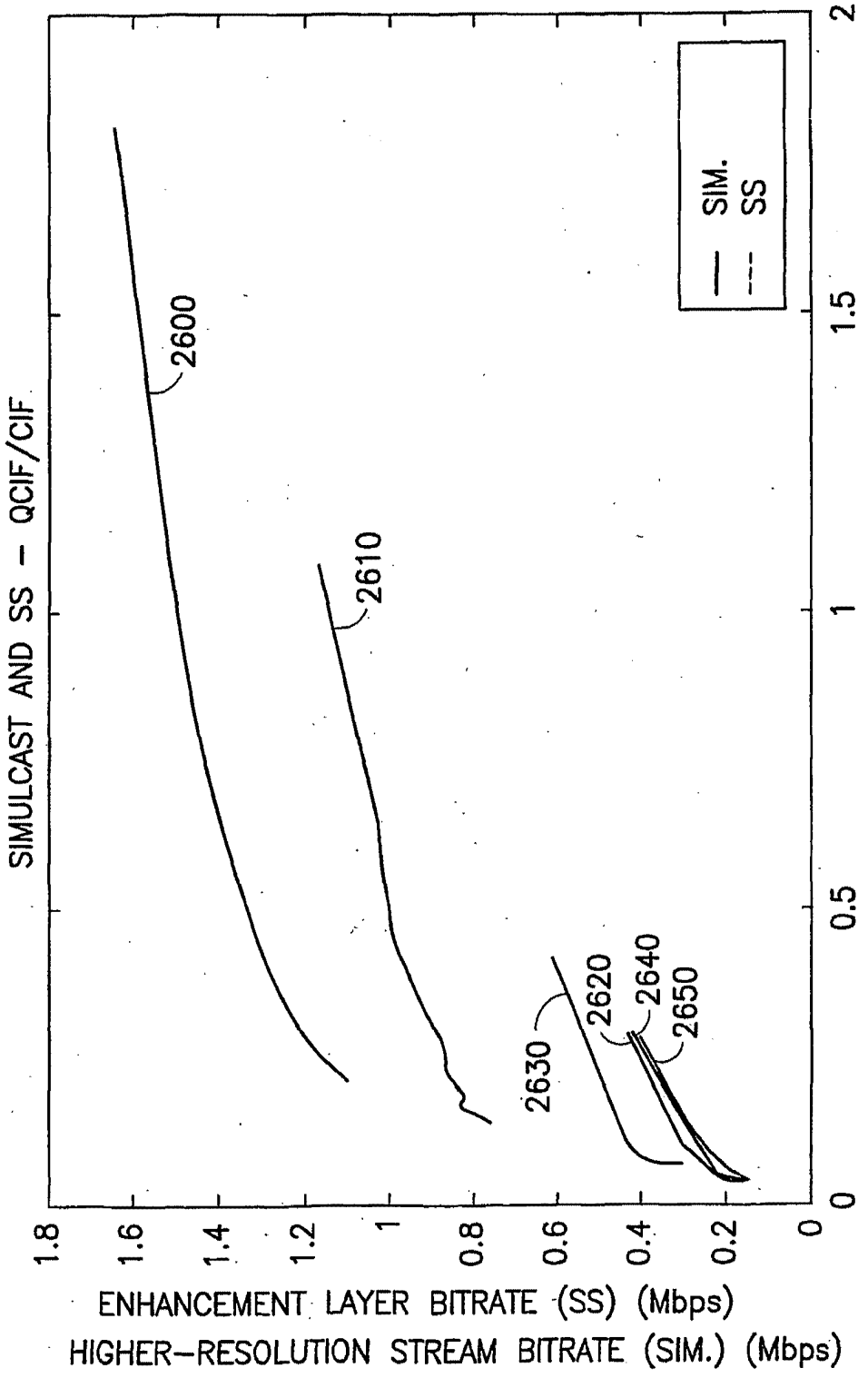

SIMULCAST AND SS - QCIF/CIF
SIM.
SS
2600
2610
2620
2630
2640
2650
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
0.5
1
1.5
2
ENHANCEMENT LAYER BITRATE (SS) (Mbps)
HIGHER-RESOLUTION STREAM BITRATE (SIM.) (Mbps)
BASE LAYER BITRATE (SS) (Mbps)
LOWER-RESOLUTION STREAM BITRATE (SIM.) (Mbps)

FIG.26

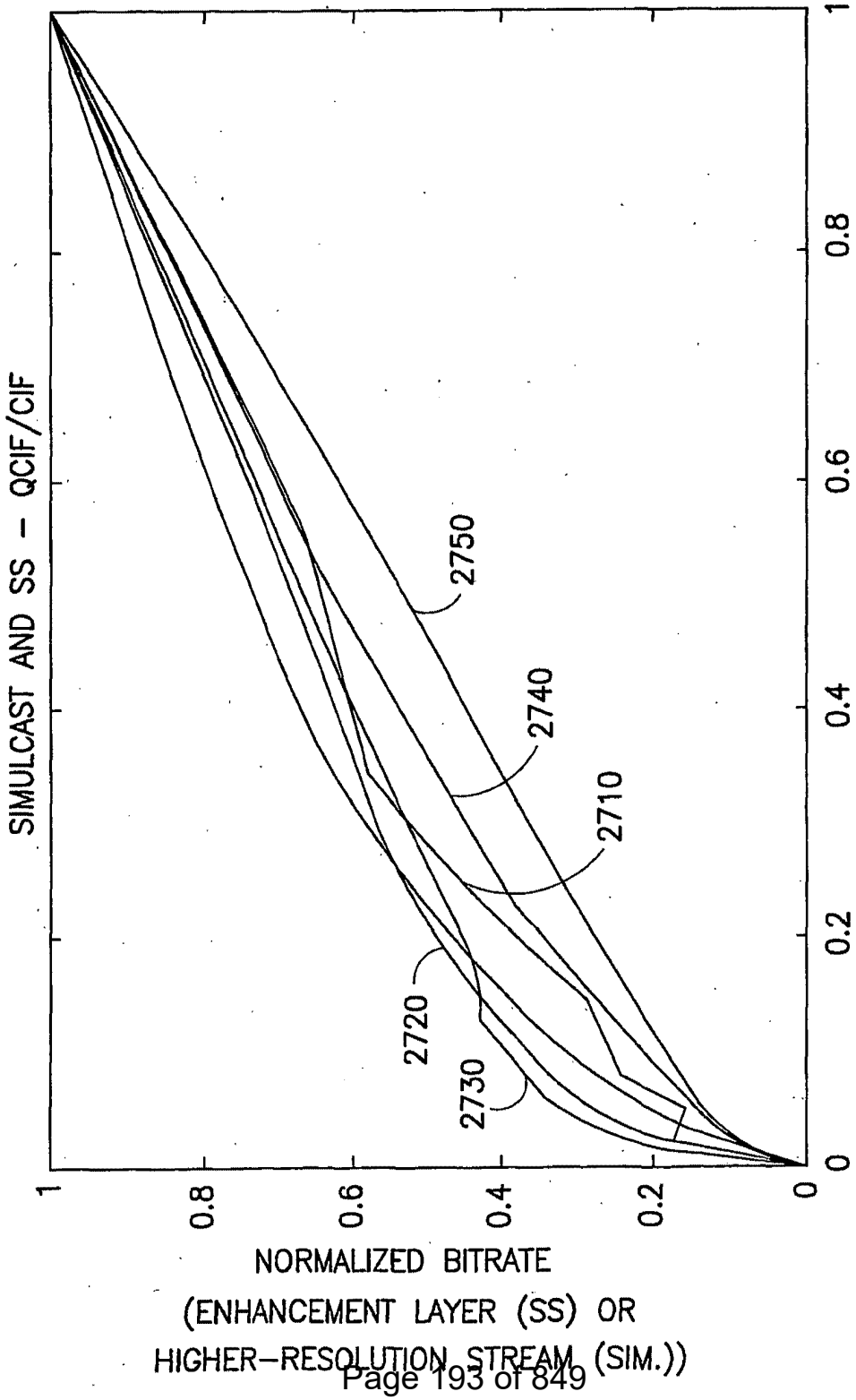

SIMULCAST AND SS – QCIF/CIF
NORMALIZED BITRATE
(ENHANCEMENT LAYER (SS) OR
HIGHER–RESOLUTION STREAM (SIM.))
1
0.8
0.6
0.4
0.2
0
2750
2740
2710
2720
2730
NORMALIZED BITRATE
(BASE LAYER (SS) OR LOWER–RESOLUTION STREAM (SIM.))
0
0.2
0.4
0.6
0.8
1

FIG.27

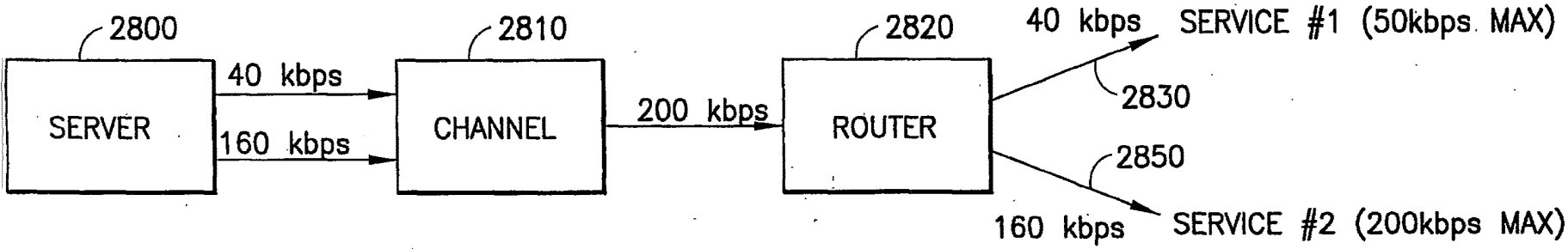

2800
SERVER
40 kbps
160 kbps
2810
CHANNEL
200 kbps
2820
ROUTER
40 kbps
SERVICE #1 (50kbps MAX)
2830
2850
160 kbps
SERVICE #2 (200kbps MAX)

FIG.28

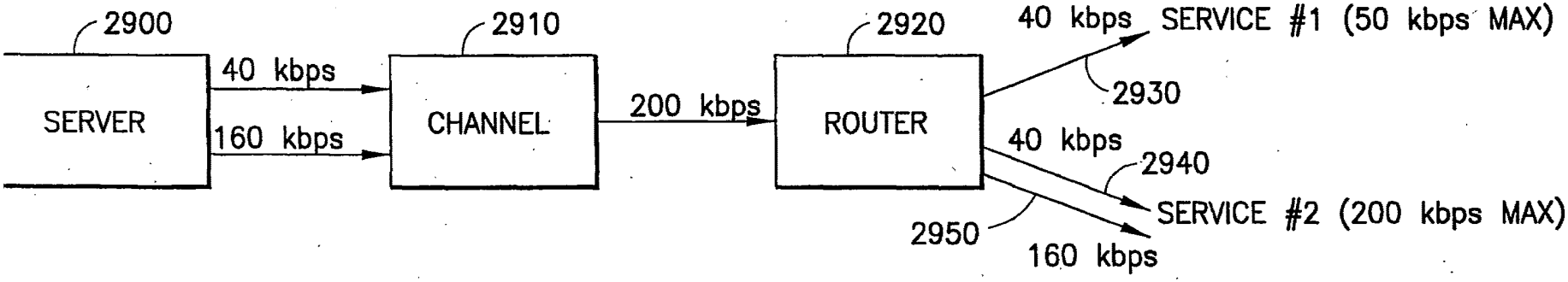

2900
SERVER
40 kbps
160 kbps
2910
CHANNEL
200 kbps
2920
ROUTER
40 kbps
SERVICE #1 (50 kbps MAX)
2930
40 kbps
2940
SERVICE #2 (200 kbps MAX)
2950
160 kbps

FIG.29

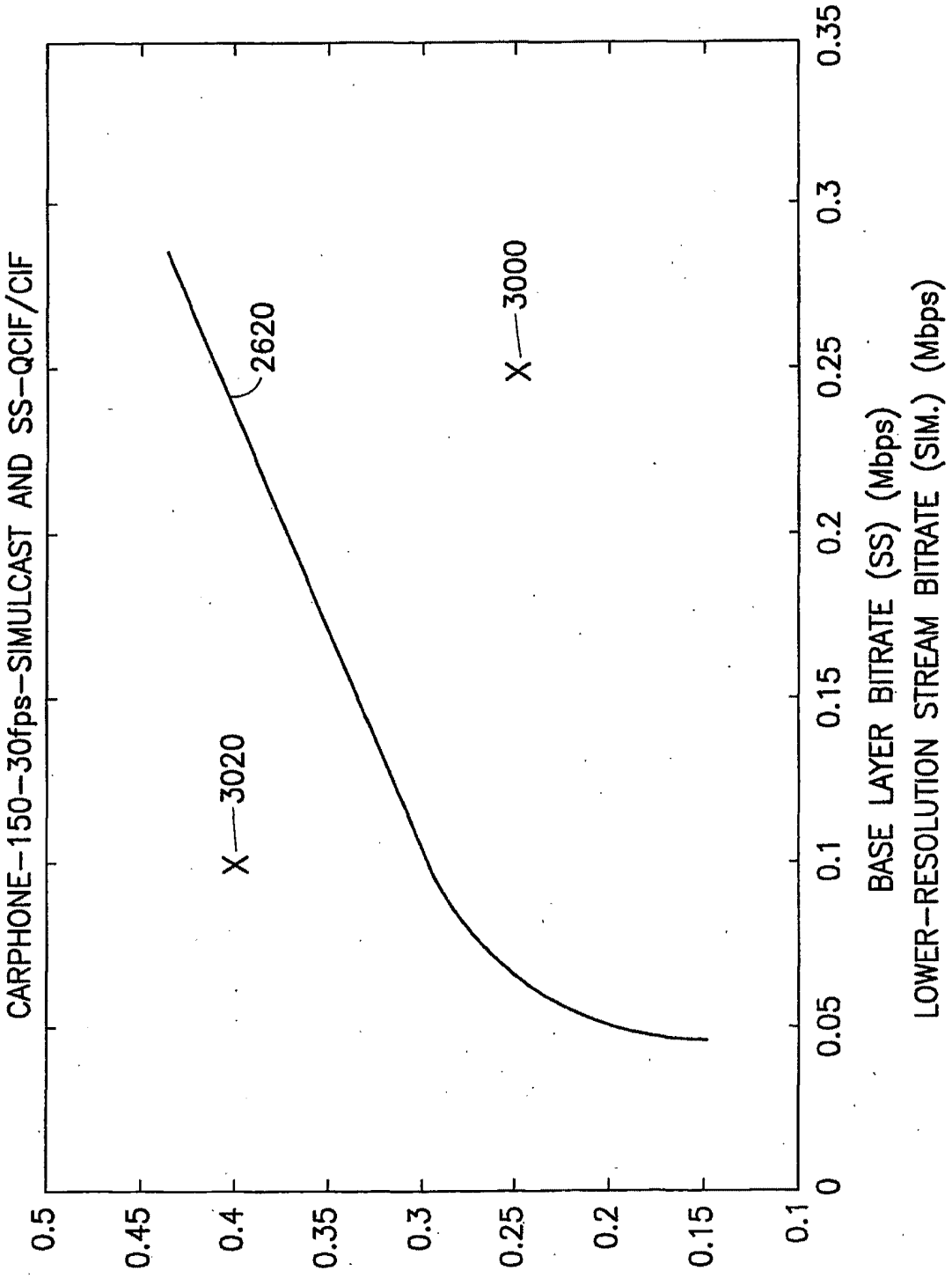

CARPHONE-150-30fps-SIMULCAST AND SS-QCIF/CIF
2620
X—3000
X—3020
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
ENHANCEMENT LAYER BITRATE (SS) (Mbps)
HIGHER-RESOLUTION STREAM BITRATE (SIM.) (Mbps)
BASE LAYER BITRATE (SS) (Mbps)
LOWER-RESOLUTION STREAM BITRATE (SIM.) (Mbps)

FIG.30

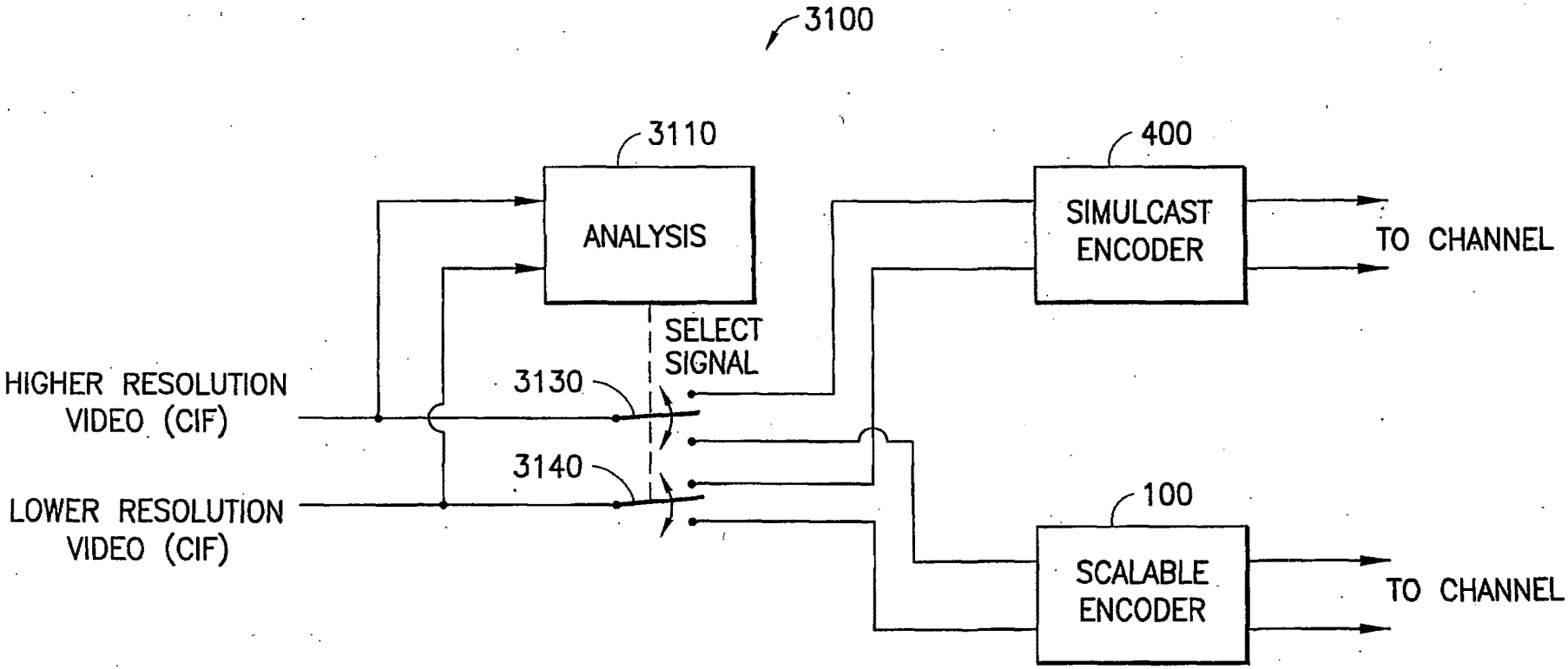

3100
3110
ANALYSIS
400
SIMULCAST ENCODER
TO CHANNEL
SELECT SIGNAL
HIGHER RESOLUTION VIDEO (CIF)
3130
3140
LOWER RESOLUTION VIDEO (CIF)
100
SCALABLE ENCODER
TO CHANNEL

FIG.31

INTERNATIONAL SEARCH REPORT

Int onal Application No
PCT/US 00/09584

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 H04N7/26

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 H04N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, WPI Data, INSPEC

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	RADHA H ET AL: "Scalable Internet video using MPEG-4" SIGNAL PROCESSING. IMAGE COMMUNICATION,NL,ELSEVIER SCIENCE PUBLISHERS, AMSTERDAM, vol. 15, no. 1-2, September 1999 (1999-09), pages 95-126, XP004180640 ISSN: 0923-5965	1-6,9, 11,34
A	abstract page 98 -page 100 page 110 -page 111 figures 1,11	30-33,40

-/--

[X] Further documents are listed in the continuation of box C. [ ] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
25 September 2000	04/10/2000

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Bergel-Faucheux, C

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
X	ILLGNER K ET AL: "SPATIALLY SCALABLE VIDEO COMPRESSION EMPLOYING RESOLUTION PYRAMIDS" IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS,US,IEEE INC. NEW YORK, vol. 15, no. 9, 1 December 1997 (1997-12-01), pages 1688-1703, XP000726010 ISSN: 0733-8716	12-17,35
A	abstract page 1688 -page 1689 page 1697 figures 6-9	30-33,40
A	CHIANG T ET AL: "HIERARCHICAL HDTV/SDTV COMPATIBLE CODING USING KALMAN STATISTICAL FILTERING" IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY,US,IEEE INC. NEW YORK, vol. 9, no. 3, April 1999 (1999-04), pages 424-437, XP000824567 ISSN: 1051-8215 abstract figures 10-17	12-40

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
11 July 2002 (11.07.2002)

PCT

(10) International Publication Number
WO 02/054776 A1

(51) International Patent Classification[7]: H04N 7/26, 7/58

(21) International Application Number: PCT/FI02/00004

(22) International Filing Date: 3 January 2002 (03.01.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/259,529	3 January 2001 (03.01.2001)	US
09/827,796	6 April 2001 (06.04.2001)	US
09/883,887	18 June 2001 (18.06.2001)	US
09/925,769	9 August 2001 (09.08.2001)	US

(71) Applicant *(for all designated States except US)*: NOKIA CORPORATION [FI/FI]; Keilalahdentie 4, FIN-02150 Espoo (FI).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: KARCZEWICZ, Marta [PL/US]; 1224 Hidden Ridge, #3097, Irving, TX 75038 (US). KURCEREN, Ragip [TR/US]; 7914 N. MacArthur Blvd., Apt. 2043, Irving, TX 75063 (US).

(74) Agent: TAMPEREEN PATENTTITOIMISTO OY; Hermiankatu 6, FIN-33720 Tampere (FI).

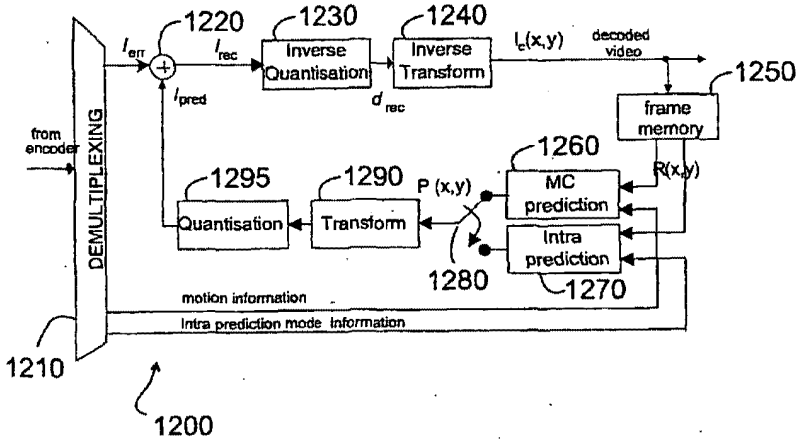
(81) Designated States *(national)*: AE, AG, AL, AM, AT, AT (utility model), AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, CZ (utility model), DE, DE (utility model), DK, DK (utility model), DM, DZ, EC, EE, EE (utility model), ES, FI, FI (utility model), GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SK (utility model), SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

[Continued on next page]

(54) Title: SWITCHING BETWEEN BIT-STREAMS IN VIDEO TRANSMISSION



(57) Abstract: The invention relates to a method for transmitting video information, in which at least a first bit-stream (510) and a second bit-stream are formed. The first bit-stream (510) comprises at least one video frame, and the second bit-stream (520) comprises at least one predictive video frame (524). At least partly different encoding parameters are used with encoding of the frames of said first bit-stream (510) and said second bit-stream (520). At least one frame of said first bit-stream (510) is being transmitted, and the transmission is switched over from said first (510) to said second bit-stream (520). In switching the transmission over from said first (510) to said second bit-stream (520), a secondary switching frame (550) is transmitted, which have been encoded using the encoding parameters of the second bit-stream (520) and at least one reference frame from the first bit-stream (510). Said secondary switching frame (550) is used as a reference frame in the reconstruction of said at least one predictive video frame (524) of the second set of video frames. The invention also relates to an encoder for encoding video information, a decoder for decoding a video information, and a signal representing encoded video information.

WO 02/054776 A1

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

SWITCHING BETWEEN BIT-STREAMS IN VIDEO TRANSMISSION

The present invention relates to a method for transmitting video information, in which at least a first bit-stream and a second bit-stream are formed from the video information, the first bit-stream comprises a first set of frames comprising at least one video frame, and the second bit-stream comprises a second set of frames comprising at least one predictive video frame, at least partly different encoding parameters are used with encoding of the frames of said first bit-stream and said second bit-stream, at least one frame of said first bit-stream is being transmitted, and the transmission is switched over from said first bit-stream to said second bit-stream. The invention also relates to an encoder comprising means for forming at least a first bit-stream and a second bit-stream from a video information, the first bit-stream comprises a first set of frames comprising at lest one video frame, and the second bit-stream comprises a second set of frames comprising at least one predictive video frame, and means for using at least partly different encoding parameters with encoding of the frames of said first bit-stream and said second bit-stream, means for transmitting at least one frame of said first bit-stream, and means for switching the transmission over from said first bit-stream to said second bit-stream. The invention further relates to a decoder for decoding a video information from a signal comprising frames from at least a first bit-stream and a second bit-stream, which are formed from the video information, the first bit-stream comprises a first set of frames comprising at lest one video frame, and the second bit-stream comprises a second set of frames comprising at least one predictive video frame, and at least partly different encoding parameters are used with encoding of the frames of said first bit-stream and said second bit-stream. The invention further relates to a signal representing encoded video information, and comprising frames from at least a first bit-stream and a second bit-stream, which are formed from the video information, the first bit-stream comprises a first set of frames comprising at lest one video frame, and the second bit-stream comprises a second set of frames comprising at least one predictive video frame, at least partly

CONFIRMATION COPY

different encoding parameters are used with encoding of the frames of said first bit-stream and said second bit-stream.

BACKGROUND

Recently, multimedia applications that include streaming audio and video information have come into greater use. Several international standardization organizations have established and proposed standards for compressing/encoding and decompressing/decoding the audio and video information. MPEG standards, established by the Motion Picture Expert Group, are the most widely accepted international standards in the field of the multimedia applications. VCEG is the "Video Coding Experts Group", working under the direction of the ITU Telecommunication Standardization Sector (ITU-T). This group works on the standard H.26L for the coding of moving pictures.

A typical video stream comprises a sequence of pictures, often referred to as frames. The frames comprise pixels arranged into a rectangular form. In existing video coding standards, such as H.261, H.262, H.263, H.26L and MPEG-4, three main types of pictures are defined: Intra frames (I-frames), Predictive frames (P-frames) and Bi-directional frames (B-frames). Each picture type exploits a different type of redundancy in a sequence of images and consequently results in different level of compression efficiency and, as explained in the following, provides different functionality within the encoded video sequence. An intra frame is a frame of video data that is coded by exploiting only the spatial correlation of the pixels within the frame itself without using any information from the past or the future frames. Intra frames are used as the basis for decoding/decompression of other frames and provide access points to the coded sequence where decoding can begin.

A predictive frame is a frame that is encoded/compressed using motion compensated prediction from a so-called reference frame, *i.e.* one or more previous/subsequent Intra frames or Predictive frames available in an encoder or in a decoder. A bi-directional frame is a frame that is

encoded/compressed by prediction from a previous Intra frame or Predictive frame and/or a subsequent Intra frame or Predictive frame.

Since adjacent frames in a typical video sequence are highly correlated, higher compression can be achieved when using Bi-directional or Predictive frames instead of Intra frames. On the other hand, when the temporal predictive coding is employed within the coded video stream, B-frames and/or P-frames cannot be decoded without correctly decoding all the other previous and/or subsequent reference frames which were used with coding of the Bi-directional and Predictive frames. In situations in which the reference frame(s) used in the encoder and respective reference frame(s) in the decoder are not identical either due to errors during transmission or due to some intentional action on the transmitting side, the subsequent frames that make use of prediction from such a reference frame can not be reconstructed on the decoding side to yield a decoded frame identical to that originally encoded on the encoding side. This mismatch is not only confined to a single frame but further propagates in time due to the use of motion compensated coding.

Figs. 1A-1C illustrate the types of encoded/compressed video frames used in a typical video encoding/decoding system. Advantageously, prior to encoding, the pictures of the video sequence are represented by these matrices of multiple-bit numbers, one representing the luminance (brightness) of the image pixels, and the other two each representing a respective one of two chrominance (colour) components. Fig. 1A depicts the way in which an Intra frame 200 is encoded using only image information present in the frame itself. Fig. 1B illustrates construction of a Predictive frame 210. Arrow 205a represents the use of motion compensated prediction to create the P-frame 210. Fig. 1C depicts construction of Bi-directional frames 220. B-frames are usually inserted between I-frames or P-frames. Fig. 2 represents a group of pictures in display order and illustrates how B-frames inserted between I-and P-frames, as well as showing the direction in which motion compensation information flows. In Figures 1B, 1C and 2, arrows 205a depict forward motion compensation

prediction information necessary to reconstruct P-frames 210, whereas arrows 215a and 215b depict motion compensation information used in reconstructing B-frames 220 in forward direction (215a) and backward direction (215b). In other words, the arrows 205a and 215a show the flow of information when predictive frames are predicted from frames that are earlier in display order than the frame being reconstructed, and arrows 215b show the flow of information when predictive frames are predicted from frames that are later in display order than the frame being reconstructed.

In motion compensated prediction, the similarity between successive frames in a video sequence is utilized to improve coding efficiency. More specifically, so-called motion vectors are used to describe the way in which pixels or regions of pixels move between successive frames of the sequence. The motion vectors provide offset values and error data that refer to a past or a future frame of video data having decoded pixel values that may be used with the error data to compress/encode or decompress/decode a given frame of video data.

The capability to decode/decompress P-frames requires the availability of the previous I- or P-reference frame, furthermore in order to decode a B-frame requires the availability of the subsequent I- or P-reference frame is also required. For example, if an encoded/compressed data stream has the following frame sequence or display order:

$I_1\ B_2\ B_3\ P_4\ B_5\ P_6\ B_7\ P_8\ B_9\ B_{10}\ P_{11} \ldots P_{n-3}\ B_{n-2}\ P_{n-1}\ I_n,$

the corresponding decoding order is:

$I_1\ P_4\ B_2\ B_3\ P_6\ B_5\ P_8\ B_7\ P_{11}\ B_9\ B_{10} \ldots P_{n-1}\ B_{n-2}\ I_n.$

The decoding order differs from the display order because the B-frames require future I- or P-frames for their decoding. Fig. 2 displays the beginning of the above frame sequence and can be referred to in order to understand the dependencies of the frames, as described earlier. P-frames require the previous I- or P-reference frame be

available. For example, P_4 requires I_1 to be decoded. Similarly, frame P_6 requires that P_4 be available in order to decode/decompress frame P_6. B-frames, such as frame B_3, require a past and/or a future I- or P-reference frame, such as P_4 and I_1 in order to be decoded. B-frames are frames between I-or P-frames during encoding.

A prior art system for encoding and decoding is shown in Figures 3 and 4. Referring to the encoder 300 of Figure 3, the frame 301 being coded I(x,y), called the current frame, is partitioned into rectangular regions of KxL pixels. The coordinates (x,y) denote the location of the pixels within the frame. Each block is encoded using either intra coding (i.e. using only the spatial correlation of the image data within the block) or inter coding (i.e. utilising both spatial and temporal prediction). The following description considers the process by which inter-coded blocks are formed. Each inter-coded block is predicted 360 from one of the previously (or subsequently) coded and transmitted frames R(x,y) in the frame memory 350, called a reference frame. The motion information used for the prediction is obtained from motion estimation and coding block 370 by using the reference frame and the current frame 305. The motion information is represented by a two-dimensional motion vector (Δx, Δy) where Δx is a horizontal and Δy is a vertical displacement. In the motion compensated (MC) prediction block, the motion vectors are used together with the reference frame to construct prediction frame P(x,y):

$$P(x,y) = R(x+\Delta x, y+\Delta y).$$

Subsequently the prediction error E(x,y), i.e. the difference between the current frame and the prediction frame P(x,y), is calculated 307 according to

$$E(x,y)= I(x,y)-P(x,y).$$

In transform block 310, the prediction error for each KxL block is represented as weighted sum of a transform basis functions $f_{ij}(x,y)$,

$$E(x, y) = \sum_{i=1}^{K} \sum_{j=1}^{L} c_{err}(i,j)\, f_{ij}(x, y).$$

The weights $c_{err}(i,j)$ corresponding to the basis functions are called transform coefficients. These coefficients are subsequently quantised in quantisation block 320 to give

$$I_{err}(i,j)=Q(c_{err}(i,j),QP)$$

where $I_{err}(i,j)$ are the quantised transform coefficients. The quantisation operation Q() introduces loss of information, but the quantised coefficients can be represented with smaller number of bits. The level of compression (loss of information) is controlled by adjusting the value of the quantisation parameter QP.

Before the motion vectors and quantised transform coefficients are supplied to the multiplexer 380, they are further encoded using Variable Length Codes (VLC). This reduces the number of bits needed to represent the motion vectors and quantised transform coefficients. The encoded motion vectors, quantised transform coefficients as well as other additional information needed to represent each coded frame is multiplexed in the multiplexer 380, and the resulting bit-stream is transmitted 415 to the decoder. The quantised transform coefficients are also forwarded to the inverse quantisation block 330 to obtain inverse quantised transform coefficients and further to the inverse transform block 340 to obtain the prediction error information $E_c(x,y)$ for the current frame. The prediction error information $E_c(x,y)$ is added to the predicted frame P(x,y) in a summing element to obtain a video frame that can be subsequently stored in the frame memory 350.

In the following, decoding video frames will be described with reference to Figure 4. The decoder 400 receives the multiplexed video bit-stream 415 from the encoder, and the demultiplexer 410 demultiplexes the bit-stream to obtain the constituent parts of the video frames to be decoded. These parts comprise at least the coded quantised prediction

error transform coefficients and the coded motion vectors, which are subsequently decoded (not shown) to obtain quantised prediction error transform coefficients and motion vectors. The quantised transform coefficients are inverse quantised in inverse quantisation block 420 to obtain inverse quantised transform coefficients $d_{err}(i,j)$ according to the following relation:

$$d_{err}(i,j)=Q^{-1}(I_{err}(i,j), QP).$$

In inverse transform block 430, the inverse quantised transform coefficients are subjected to an inverse transform to obtain prediction error $E_c(x,y)$:

$$E_c(x,y) = \sum_{i=1}^{K} \sum_{j=1}^{L} c_{err}(i,j)\, f_{ij}\,(x,y).$$

The pixels of the current coded frame are reconstructed by finding the prediction pixels in the reference frame R(x,y) obtained from frame memory 440, and using the received motion vectors together with the reference frame in motion compensation prediction block 450, to obtain a predicted frame P(x,y). The predicted frame P(x,y) and the prediction error information $E_c(x,y)$ are summed together in a summing element 435 according to the following relation:

$$I_c(x,y)= R(x+\Delta x, y+\Delta y)+E_c(x,y).$$

These values $I_c(x,y)$ can further be filtered to obtain decoded video frames 445. The values $I_c(x,y)$ are also stored in the frame buffer 440. The reconstructed values $I_c(x,y)$ can be filtered in a filtering block (not shown in Figure 4) after the summation block 435.

Video streaming has emerged as an important application in the fixed Internet. It is further anticipated that video streaming will also be important in and in the near future in 3G wireless networks. In streaming applications the transmitting server starts transmitting a pre-

encoded video bit stream via a transmission network to a receiver upon a request from the receiver. The receiver plays back the video stream back while receiving it. The best-effort nature of present networks causes variations in the effective bandwidth available to a user due to the changing network conditions. To accommodate these variations, the transmitting server can scale the bit rate of the compressed video. In the case of a conversational service characterized by real-time encoding and point-to-point delivery, this can be achieved by adjusting the source encoding parameters on the fly. Such adjustable parameters can be, for example, a quantisation parameter, or a frame rate. The adjustment is advantageously based on feedback from the transmission network. In typical streaming scenarios when a previously encoded video bit stream is to be transmitted to the receiver, the above solution cannot be applied.

One solution to achieve bandwidth scalability in case of pre-encoded sequences is to produce multiple and independent streams having different bit-rates and quality. The transmitting server then dynamically switches between the streams to accommodate variations in the available bandwidth. The following example illustrates this principle. Let us assume that multiple bit streams are generated independently with different encoding parameters, such as quantisation parameter, corresponding to the same video sequence. Let $\{P_{1,n-1}, P_{1,n}, P_{1,n+1}\}$ and $\{P_{2,n-1}, P_{2,n}, P_{2,n+1}\}$ denote the sequence of decoded frames from bit streams 1 and 2, respectively. Since the encoding parameters are different for the two bit streams, frames reconstructed from them at the same time instant, for example, frames $P_{1,n-1}$ and $P_{2,n-1}$, are not identical. If it is now assumed that the server initially sends encoded frames from bit stream 1 up to time *n* after which it starts sending encoded frames from bit stream 2, the decoder receives frames $\{P_{1,n-2}, P_{1,n-1}, P_{2,n}, P_{2,n+1}, P_{2,n+2}\}$. In this case $P_{2,n}$ cannot be correctly decoded since its reference frame $P_{2,n-1}$ is not received. On the other hand, the frame $P_{1,n-1}$, which is received instead of $P_{2,n-1}$, is not identical to $P_{2,n-1}$. Therefore switching between bit streams at arbitrary locations leads to visual artefacts due to the mismatch between the reference frames used for motion compensated prediction in the different sequences.

These visual artefacts are not only confined to the frame at the switching point between bit streams, but propagates in time due to the continued motion compensated coding in the remaining part of the video sequence.

In current video encoding standards, perfect (mismatch-free) switching between bit streams is only possible at positions where the current and future frames or regions thereof do not use any information previous to the current switching location, i.e., at I-frames. Furthermore, by placing I-frames at fixed (e.g. 1 sec) intervals, VCR functionalities, such as random access or "Fast Forward" and "Fast Backward" (increased playback rate) for streaming video content, are achieved. The user may skip a portion of the video sequence and restart playing at any I-frame location. Similarly, an increased playback rate can be achieved by transmitting only I–frames. The drawback of using I-frames in these applications is that since I-frames do not exploit any temporal redundancy they require a much larger number of bits than P-frames at the same quality.

SUMMARY OF THE INVENTION

It is an aim of the present invention to provide a new method and a system for transmitting video images in variable transmission environments. The invention is based on the idea that correct (mismatch-free) switching between video streams is enabled by forming a new type of a compressed video frame and inserting frames of the new type into video bit-streams at locations where switching from one bit-stream to another is to be allowed. In this description, the new type of compressed video frame will be referred to generally as an S-frame. More specifically, S-frames may be classified as SP-frames, which are formed at the decoder using motion compensated prediction from already decoded frames using motion vector information, and SI-frames, which are formed at the decoder using spatial (intra) prediction from already decoded neighbouring pixels within a frame being decoded. In general, an S-frame according to the invention is formed on a block-by-block basis and may comprise both inter-coded (SP)

blocks as well as intra-coded (SI) blocks. The method according to the present invention is primarily characterized in that in switching the transmission over from said first bit-stream to said second bit-stream, the secondary bit-stream comprises at least one primary switching frame, a secondary switching frame is transmitted, the secondary switching frame having been encoded using the encoding parameters of the second bit-stream and at least one reference frame from the first bit-stream, and said secondary switching frame is used in place of the primary switching frame as a reference frame to be used in the reconstruction of said at least one predictive video frame of the second set of video frames. The encoder according to the present invention is primarily characterized in that said means for switching the transmission over from said first bit-stream to said second bit-stream comprise means for encoding a secondary switching frame using the encoding parameters of the second bit-stream and reference frames from the first bit-stream to enable to switch the transmission over from said first bit-stream to said second bit-stream. The decoder according to the present invention is primarily characterized in that the decoder comprises means for decoding a secondary switching frame, said secondary switching frame having been encoded using the encoding parameters of the second bit-stream and at least one reference frame from the first bit-stream, and added to said signal in place of a primary switching frame as a reference frame to be used in the reconstruction of said at least one predictive video frame of the second set of video frames, and that said means for decoding a secondary switching frame comprise means for using decoding parameters of the second bit-stream and reference frames from the first bit-stream. The signal according to the present invention is primarily characterized in that in switching the transmission over from said first bit-stream to said second bit-stream, the secondary bit-stream comprises at least one primary switching frame, said signal comprises a secondary switching frame having been encoded using the encoding parameters of the second bit-stream and at least one reference frame from the first bit-stream, and said secondary switching frame is in place of the primary switching frame as a reference frame to be used in the reconstruction

of said at least one predictive video frame of the second set of video frames.

Considerable advantages are achieved by the present invention compared to the methods and systems of prior art. The invention allows switching between bit streams to occur not only at the locations of I-frames but also at the locations of the SP-frames. The coding efficiency of an SP-frame is much better than the coding efficiency of a typical I-frame wherein less bandwidth is needed to transmit bit streams having SP-frames in locations where, according to prior art, I-frames would be used, while still providing enough adaptability to changing transmission conditions. The switching of one bit stream into another can be performed at locations in which an SP-frame according to the invention is placed in the encoded bit stream. The images reconstructed from the bit stream by the decoder is not degraded as a result of changing from one bit stream to another. The invention also has the advantage that random access, fast-forward and fast rewind operations can be performed on the bit stream. The system according to the invention offers improved error recovery and resiliency properties compared with the described prior art solutions.

These and other features, aspects, and advantages of embodiments of the present invention will become apparent with reference to the following description in conjunction with the accompanying drawings. It is to be understood, however, that the drawings are designed solely for the purposes of illustration and not as a definition of the limits of the invention, for which reference should be made to the appended claims.

BRIEF DESCRIPTION OF THE DRAWINGS

In the following the invention will be described in more detail with reference to the appended drawings, in which

Figs. 1A-1C and 2 are diagrams showing the prior art encoding/compression of video frames,

Fig. 3 is a block diagram of a generic motion-compensated predictive video coding system (encoder),

Fig. 4 is a block diagram of a generic motion-compensated predictive video coding system (decoder),

Fig. 5 is an illustration showing switching between two different bit streams using S-frames according to the invention,

Fig. 6 is a block diagram of a decoder in accordance with an advantageous embodiment of the invention,

Fig. 7 is an illustration of random access using S-frames,

Fig. 8 is an illustration of a fast-forward process using S-frames,

Fig. 9 is a block diagram of a decoder in accordance with another advantageous embodiment of the invention,

Fig. 10 is a block diagram of an decoder in accordance with yet another advantageous embodiment of the invention,

Fig. 11 is a block diagram of an encoder in accordance with an advantageous embodiment of the invention,

Fig. 12 is a block diagram of a decoder in accordance with another advantageous embodiment of the invention,

Fig. 13 is a block diagram of an encoder according to a second embodiment of the invention,

Fig. 14 is an illustration of a error resiliency/recovery process using SP-frames,

Fig. 15 is a block diagram of an encoder in accordance with the third advantageous embodiment of the invention,

Fig. 16 is a block diagram of a decoder in accordance with yet another advantageous embodiment of the invention.

DETAILED DESCRIPTION OF THE INVENTION

In the following the invention is described in a system in which multiple bit streams are formed from a video signal. The video signal can be any digital video signal comprising multiple images, *i.e.* an image sequence. The digital video signal is encoded in an encoder to form multiple bit streams. Each bit stream is encoded from the same video signal using at least partly different encoding parameters. For example, the bit rate can be altered by selecting the encoding parameters differently, and in this way bit streams with different bit rates can be formed. The encoding parameters can be, for example, frame rate, quantisation parameter, spatial resolution, or another factor affecting the images size, known as such to a person skilled in the art. The encoder also inserts at least one Intra frame to each bit stream. Typically, at least the first frame of each bit stream is preferably an Intra frame. This enables the decoder to start reconstruction of the video signal. The encoder used in the encoding of the I-frames, P-frames and B-frames can be any prior art encoder which performs the encoding of the video signal, or there may exist more than one prior art encoder each using different encoding parameters to form multiple bit streams. However, to encode a video signal containing also SP-frames and/or SI-frames according to the invention, new functionality is needed in the encoder. This will be explained later in greater detail.

The encoder also inserts frames encoded using motion compensated predictive coding (P-frames and optionally B-frames) into the bit streams. The encoder also inserts a new type of frames, referred to as S-frames in this description, into each bit stream at locations where switching between different bit streams will be allowed according to the invention. The S-frames may be used at locations where in prior art methods an Intra coded frame would be inserted, or the S-frames may be used in addition to using Intra coded frames in the video sequence.

Advantageously, the different bit streams are stored into storage means for later use. However, it is also possible that transmission may take place substantially immediately after encoding wherein it is not necessary to store complete video sequences, but storing the necessary reference frames suffices. Transmission of the encoded video stream may be performed *e.g.* by a transmitting server, which has means for retrieving the stored bit streams for transmission, and/or means for receiving the bit streams directly from the encoder. The transmitting server also has means for transmitting the bit stream to the transmission network (not shown).

In the following, a method according to an advantageous embodiment of the invention will be described. Fig. 5 depicts part of a first bit stream 510 and part of a second bit stream 520, which are formed in the encoder. Only a few P-frames of the respective bit streams are shown. Specifically, the first bit stream 510 is shown to comprise P-frames 511, 512, 514 and 515, while the second bit stream 520 comprises corresponding P-frames 521, 522, 524 and 525. Both the first bit stream 510 and the second bit stream 520 also comprise S-frames 513 (also labelled S_1), 523 (also labelled S_2) at corresponding locations. It is assumed that the two bit streams 510 and 520 correspond to the same sequence encoded at different bit rates for example by using different frame rates, different spatial resolutions or different quantisation parameters. It is further assumed that the first bit stream 510 is being transmitted from the transmitting server to a decoder 600, 1200, 1300 (Figs. 6, 12 and 13, respectively) via a transmission network, and that the transmitting server receives a request from the transmission network to change the bit rate of the video stream being transmitted.

As mentioned above, S-frames are placed in the bit stream during the encoding process at those locations within the video sequences where switching from one bit stream to another is allowed. As can be seen from Figure 5, in an advantageous embodiment of the invention a further S-frame 550 (also labelled S_{12}) is associated with S-frames S_1 and S_2. This S-frame is called a secondary representation of the S-

frame (or secondary S-frame for short), and is only transmitted during bit stream switching. This secondary S-frame S_{12} is generated by a special encoding of the nth frame in the video sequence using the encoding parameters of the corresponding S-frame 523 (S_2) in the second bit stream 520 and using motion compensated prediction from the reference frames of the nth frame in the first bit stream 510. It should be noted that in the situation shown in Figure 5, S-frame S_2 uses the previously reconstructed frames from the second bit stream 520 as reference frames and the secondary S-frame S_{12} uses the previously reconstructed frames from the first bit stream 510 as reference frames. However, the reconstructed pixel values of both S_2 and S_{12} are identical. S-frame S_{12} is transmitted only when switching from the first bit stream 510 to the second bit stream 520 is actually performed. Therefore it is not necessary to form the secondary S-frames during the encoding phase but only when switching is performed. On the other hand, it may be useful to have at least some secondary S-frames formed earlier at the time when the different bit streams are formed to reduce computational load during transmission.

When the transmitting server reaches the frame of the video sequence encoded as S-frame 513 (S_1) in the first bit stream 510, it can begin the necessary operations to continue transmission of the video stream using the encoded frames of the second bit stream 520. At that point the transmitting server has already transmitted P-frames 511 and 512 from the first bit stream 510 and the decoder 600, 1200, 1300 has received and decoded the respective P-frames 511, 512. Thus, those frames have already been stored in the frame memory 640, 1250, 1360 of the decoder 600, 1200, 1300. The frame memory 640, 1250, 1360 comprises sufficient memory to store all those frames, which are needed to reconstruct a P-frame or a B-frame, *i.e.* the necessary information of all the reference frames required by the current frame to be reconstructed.

The transmitting server performs the following operations to continue the transmission of the video stream using the encoded frames of the second bit stream 520. The transmitting server notices, for example, by

examining the type information of the frame, that the current frame to be transmitted is an S-frame, so it is possible to perform switching between the bit streams. Of course, switching is only performed if a request to do so has been received or there is for some other reason a need to perform the switch. The transmitting server inputs the corresponding S-frame 523 of the second bit stream, uses it to form the secondary S-frame 550 (S_{12}), and transmits the secondary S-frame (S_{12}) to the decoder 600, 1200, 1300. The transmitting server does not transmit the S-frame S_2 of the second bit stream, but sends the secondary S-frame S_{12} in its place. The secondary S-frame is formed in such a way that by decoding the secondary S-frame S_{12} the decoder 600 can reconstruct an image identical to that which would be produced if it were using respective frames 521, 522 and the S-frame 523 of the second bit stream 520. After the transmission of the secondary S-frame the transmitting server continues to transmit the encoded frames if the second bit stream 520, i.e., 524, 525, and so on.

The S-frame 513, 523, 550 can comprise blocks encoded using only spatial correlation among the pixels (intra blocks) and blocks encoded using both spatial and temporal correlation (inter blocks). For each inter block the prediction of this block, P(x,y), is formed in the decoder 600, 1200, 1300 using received motion vectors and a reference frame. The transform coefficients c_{pred} for P(x,y) corresponding to basis functions $f_{ij}(x,y)$ are calculated and quantised. The quantised values of the transform coefficients c_{pred} are denoted as l_{pred} and the dequantised values of the quantised transform coefficients l_{pred} are denoted as d_{pred}. Quantised coefficients l_{err} for the prediction error are received from the encoder. The dequantised values of these coefficients will be denoted as d_{err}. Value of each pixel S(x,y) in the inter block is decoded as a weighted sum of the basis functions $f_{ij}(x,y)$ where the weight values d_{rec} will be called dequantised reconstruction image coefficients. The values of d_{rec} have to be such that coefficients c_{rec} exist from which by quantisation and dequantisation d_{rec} can be obtained. In addition, values d_{rec} have to fulfill one of the following conditions:

$d_{rec} = d_{pred} + d_{err}$, or

$c_{rec} = c_{pred} + d_{err}$.

Values S(x,y) can further be normalized and filtered.

Next, encoding of the S-frames placed within the bit stream, e.g. S-frames 513 (S_1) and 523 (S_2) is described.

Generally, an S-frame according to the invention such as frames 513, 523 in Figure 5 is constructed on a block-by-block basis. As explained earlier, each of the blocks may be coded either in such a way as to take advantage of the spatial correlations among pixels of the image being coded (intra or SI-blocks) or in such a way as to take advantage of the temporal correlation between blocks of pixels in successive frames of a video sequence (inter or SP-blocks).

The encoding of S-frames according to the invention will be described with reference to Figure 11 which is a block diagram of an S-frame encoder 1100 according to a first embodiment of the invention.

A video frame to be encoded in S-frame format is first partitioned into blocks and each block is then encoded as either an SP-block, an SI-block, or an intra-block, an intra-block being known as such from prior art. Switch 1190 is operated as appropriate to switch between the SI and SP encoding modes, i.e., the switch 1190 is a construction used in the description of the invention, not necessarily a physical device. In SP-encoding mode switch 1190 is operated to obtain a motion compensated prediction for the current block 1170. Motion compensated prediction block 1170 forms a prediction P(x,y) for the current block of the frame being encoded in a manner analogous to that used in motion compensated prediction known from prior art. More specifically, motion compensated prediction block 1170 forms the prediction P(x,y) for the current block of the frame being encoded by determining a motion vector describing the relationship between the pixels in the current block and pixel values of a reconstructed reference frame held in frame memory 1146.

In SI-encoding mode switch 1190 is operated to obtain a prediction for the current block of the frame being coded from intra prediction block 1180. Intra prediction block 1180 forms the prediction P(x,y) for the current block of the frame being encoded in a manner analogous to that used in intra prediction known from prior art. More specifically, intra prediction block 1180 forms the prediction P(x,y) for the current block of the frame being encoded using spatial prediction from already encoded neighbouring pixels within the frame being encoded.

In both SP- and SI-coding modes the prediction P(x,y) takes the form of a block of pixel values. A forward transform, for example a Discrete Cosine Transform (DCT), is applied to the predicted block of pixel values P(x,y) in block 1160 and the resulting transform coefficients, referred to as c_{pred}, are subsequently quantised in quantisation block 1150 to form quantised transform coefficients l_{pred}. Corresponding operations are also performed on the original image data. More specifically, the current block of pixel values of the original image being encoded is applied to transform block 1110. Here, a forward transform (e.g., a DCT) is applied to the pixel values of the original image block to form transform coefficients c_{orig}. These transform coefficients are passed to quantisation block 1120 where they are quantised to form quantised transform coefficients l_{orig}. The summing element 1130 receives both sets of quantised transform coefficients l_{pred} and l_{orig} from the respective quantisation blocks 1150 and 1120 and generates a set of quantised prediction error coefficients l_{err} according to the relationship:

$$l_{err} = l_{orig} - l_{pred} .$$

The quantised prediction error coefficients l_{err} are passed to multiplexer 1135. If the current block is encoded in SP-format/mode, multiplexer 1135 also receives the motion vectors for the SP-coded block. If the current block is encoded in SI-format/mode, information concerning the intra prediction mode used to form the prediction for the SI-coded block in intra prediction block 1180 is passed to the multiplexer. Advantageously, variable length coding is applied to the quantised

prediction error coefficients l_{err} and to the motion vector or intra prediction mode information in the multiplexer 1135, a bit-stream is formed by multiplexing together the various forms of information and the bit-stream thus formed is transmitted to a corresponding decoder 1200, 1300 (see Figs. 12 and 13).

The S-frame encoder 1100 according to the invention also comprises local decoding functionality. The quantised prediction transform coefficients l_{pred} formed in quantisation block 1150 are supplied to the summing element 1140 which also receives the quantisation error coefficients l_{err} . The summing element 1140 recombines the quantised prediction transform coefficients l_{pred} and the quantised prediction error coefficients l_{err} to form a set of reconstructed quantised transform coefficients l_{rec} according to the relationship:

$$l_{rec} = l_{pred} + l_{err} .$$

The reconstructed quantised transform coefficients are passed to inverse quantisation block 1142 which inverse quantises the reconstructed quantised transform coefficients to form inverse quantised reconstructed transform coefficients d_{rec}. The inverse quantised reconstructed transform coefficients are further passed to inverse transform block 1144 where they are subjected to e.g. an Inverse Discrete Cosine Transform (IDCT), or any other inverse transform corresponding to the transform performed in block 1160. As a result, a block of reconstructed pixel values is formed for the image block in question and is stored in frame memory 1146. As subsequent blocks of the frame being encoded in S-frame format undergo the previously described encoding and local decoding operations, a decoded version of the current frame is progressively assembled in the frame memory from where it can be accessed and used in intra prediction of subsequent blocks of the same frame or in inter (motion compensated) prediction of subsequent frames in the video sequence.

Operation of a generic S-frame decoder according to a first embodiment of the invention will now be described with reference to Figure 12.

The bit-stream generated by the S-frame encoder previously described in connection with Figure 11 is received by decoder 1200 and is demultiplexed into its constituent parts by demultiplexer 1210. The decoder reconstructs a decoded version of the S-frame on a block-by-block basis. As previously described, an S-frame may comprise intra-blocks, SP-coded and SI-coded image blocks. For SP-format image blocks, the information in the received bit-stream comprises VLC encoded motion coefficient information and VLC encoded quantised prediction error coefficients I_{err}. For image blocks encoded in SI-format the information in the received bit-stream comprises VLC coded information relating to the intra prediction mode used to form the intra prediction for the SI-coded block together with VLC coded quantised prediction error coefficients I_{err}.

When decoding an SP-coded block, demultiplexer 1210 first applies appropriate variable length decoding (VLD) to the received bit-stream to recover the motion vector information and quantised prediction error coefficients I_{err}. It then separates the motion vector information from the quantised prediction error coefficients I_{err}. The motion vector information is supplied to motion compensated prediction block 1260 and the quantised prediction error coefficients recovered from the bit-stream are applied to one input of summing element 1220. The motion vector information is used in motion compensated prediction block 1260 together with pixel values of a previously reconstructed frame held in frame memory 1250 to form a prediction P(x,y) in a manner analogous to that employed in the encoder 1100.

When decoding an SI-coded block, demultiplexer 1210 applies appropriate variable length decoding to the received intra prediction mode information and the quantised prediction error coefficients I_{err}. The intra prediction mode information is then separated from the quantised prediction error coefficients and supplied to intra prediction

block 1270. The quantised prediction error coefficients l_{err} are supplied to one input of the summing element 1220. The intra prediction mode information is used in intra prediction block 1270 in conjunction with previously decoded pixel values of the current frame held in frame memory 1250 to form a prediction P(x,y) for the current block being decoded. Again, the intra prediction process performed in decoder 1200 is analogous to that performed in encoder 1100 and previously described.

Once a prediction for the current block of the frame being decoded has been formed, switch 1280 is operated so that the prediction P(x,y) which comprises predicted pixel values is supplied to transform block 1290. Again, switch 1280 is an abstract construction used in the description of the invention, not necessarily a physical device. In the case of an SP-coded block, switch 1280 is operated to connect motion compensated prediction block 1260 to transform block 1290, while in the case of an SI-coded block it is operated to connect intra prediction block 1270 to transform block 1290.

In block 1290, a forward transform, e.g., a Discrete Cosine Transform (DCT), is applied to the predicted block of pixel values P(x,y) and the resulting transform coefficients c_{pred} are supplied to quantisation block 1295 where they are quantised to form quantised transform coefficients l_{pred}. Quantised transform coefficients l_{pred} are then supplied to the second input of summing element 1220 where they are added to the prediction error coefficients l_{err} to form reconstructed quantised transform coefficients l_{rec} according to the relationship:

$$l_{rec} = l_{pred} + l_{err}.$$

The reconstructed quantised transform coefficients l_{rec} are further supplied to inverse quantisation block 1230 where they are inverse quantised to form inverse quantised reconstructed transform coefficients d_{rec}. The inverse quantised transform coefficients d_{rec} are then passed to inverse transform block 1240 where they are subjected to e.g. an Inverse Discrete Cosine Transform (IDCT), or any other

inverse transform corresponding to the transform performed in block 1290. In this way, a block of reconstructed pixel values is formed for the image block in question. The reconstructed pixel values are supplied to the video output and to frame memory 1250. As subsequent blocks of the S-frame being decoded undergo the previously described decoding operations, a decoded version of the current frame is progressively assembled in frame memory 1250 from where it can be accessed and used in the intra prediction of subsequent blocks of the same frame or in the inter (motion compensated) prediction of subsequent frames in the video sequence.

Having reviewed the structure and function of an S-frame encoder and decoder according to a first embodiment of the invention, it is now possible to understand how S-frames according to the invention make it possible to switch between bit-streams without mismatch errors such as those encountered in prior video encoding/decoding systems. Referring once more to the bit-stream switching example depicted in Figure 5, switching from the first bit-stream 510 to the second bit-stream 520 occurs at the location of S-frames S_1 (513) and S_2 (523) in the respective bit-streams. As previously noted, when switching is performed, a secondary S-frame, denoted S_{12} (550) is encoded and transmitted. The secondary S-frame S_{12} is encoded using the encoding parameters of the second bit-stream 520 and reference frames from the first bit-stream 510 in such a way that when the secondary frame S_{12} is decoded its reconstructed pixel values are identical to those that would have resulted from transmission of frame S_2 in the second bit-stream.

Let I^2_{err} and I^2_{pred} denote the quantised coefficients of the prediction error and the prediction frame, respectively, obtained from encoding of the SP-frame S_2 with the procedure described above, and let I^2_{rec} denote the quantised reconstructed image coefficients of the S-frame S_2. The encoding of the secondary S-frame 550 (S_{12}) follows the same procedures as in the encoding of the S-frame 523 (S_2) with the following exceptions: 1) The reference frame(s) used in the prediction of each block of the secondary S-frame S_{12} are the reconstructed

frames obtained by decoding the first bit stream 510 up to the current n^{th} frame in the video sequence, 2) The quantised prediction error coefficients are calculated as follows: $I^{12}_{err} = I^{2}_{rec} - I^{12}_{pred}$ where I^{12}_{pred} denotes the quantised prediction transform coefficients. The quantised prediction error coefficients I^{12}_{err} and the motion vectors are transmitted to the decoder 1200.

When decoding the secondary S-frame S_{12} in the decoder 1200 using the reconstructed frames from the first bit stream 510 before the switch as reference frames, coefficients I^{12}_{pred} of the secondary S-frame are constructed and added to the received quantised prediction error coefficients I^{12}_{err} as described above, i.e., $I^{12}_{rec} = I^{12}_{err} + I^{12}_{pred} = I^{2}_{rec} - I^{12}_{pred} + I^{12}_{pred} = I^{2}_{rec}$. It can be noted from this equation that I^{12}_{rec} and I^{2}_{rec} are identical. Therefore, although the secondary S-frame S_{12} and the S-frame of the second bit stream S_2 have different reference frames, when S_{12} is decoded it produces an image with reconstructed pixel values identical to those that would result from decoding S-frame S_2.

From the foregoing description of the encoding and decoding of S-frames according to the invention, it will be appreciated that a significant difference exists compared with encoding and decoding of P-frames and I-frames according to prior art, specifically, it should be appreciated that when encoding or decoding an image block in SP or SI -format, the prediction P(x,y) for the block in question is transformed into the transform coefficient domain by applying a transformation such as a Discrete Cosine Transform. The transform coefficients thus produced are subsequently quantised and a prediction error is determined in the quantised coefficient domain. This contrasts with predictive coding according to prior art in which a prediction error is determined in the spatial (pixel value) domain.

In the following, operation of the decoder 1200 during switching between bit-streams 510 and 520 is described in detail. At the position in the video sequence where switching from the first bit-stream to the second bit-stream 520 takes place, the decoder 1200 has already

received and decoded the previous P-frames 511 and 512 of the first bit-stream 510. The decoded frames are stored in frame memory 1250 and are thus available for use as reference frames. When switching from the first bit-stream 510 to the second bit-stream 520 takes place, encoder 1100 (figure 11) constructs and encodes secondary S-frame S_{12} (550) and transmits the encoded video information representing S_{12} to decoder 1200.

As previously explained, encoding is performed on a block-by-block basis. Specifically, secondary S frame S_{12} is encoded as an assembly of image blocks and generally each image block is encoded as either an SP-coded block or an SI-coded block or an intra-block. For the SP-coded blocks of the secondary S-frame S_{12}, the compressed video information transmitted from encoder to decoder takes the form of quantised prediction error transform coefficients I^{12}_{err} and motion vector information. For the SI-coded blocks of the secondary S-frame S_{12} the compressed video information comprises quantised prediction error transform coefficients I^{12}_{err} and information relating to the intra prediction mode used in forming the prediction for the SI-coded block in the encoder. As previously explained, the compressed video information is subjected to appropriate variable length coding (VLC) prior to transmission from the encoder in order to further reduce the number of bits required for its representation.

The compressed video information for a given image block is received at decoder 1200 and is first subjected to appropriate variable length decoding (VLD) and is separated into its constituent parts by demultiplexer 1210. The quantised prediction error coefficients I^{12}_{err} extracted from the received bit-stream are applied to the first input of summer 1220 and a block of predicted pixel values P(x,y) is formed for each image block according to its coding mode (SP or SI). In the case of an SP-coded block, the block of predicted pixel values P(x,y) is formed in motion compensated prediction block 1260 using a reference frame from the first bit-stream (e.g. P-frame 511 or 512) available in frame memory 1250 and the motion vector information extracted from the encoded video information of secondary S-frame S_{12} by

demultiplexer 1210. In the case of an SI-coded block, a block of predicted pixel values P(x,y) is formed in intra prediction block 1270 using previously decoded pixels of the secondary S-frame S_{12} also stored in frame memory 1250. Intra prediction is performed according to the intra prediction mode information extracted from the received video information for secondary S-frame S_{12} by demultiplexer 1210.

Once a prediction for the current block of the secondary S-frame has been formed the predicted pixel values P(x,y) are passed to transform block 1290. Here a forward transform (e.g. a Discrete Cosine Transform (DCT)) is applied the predicted pixel values P(x,y) to form a set of transform coefficients c_{pred}. These transform coefficients are then passed to quantisation block 1295 where they are quantised to form quantised transform coefficients l^{12}_{pred}. The quantised transform coefficients l^{12}_{pred} are, in turn, applied to the second input of summer 1220. Summer 1220 combines the quantised transform coefficients l^{12}_{pred} with quantised prediction error transform coefficients l^{12}_{err} to form reconstructed quantised transform coefficients l^{12}_{rec} according to the relationship:

$$l^{12}_{rec} = l^{12}_{pred} + l^{12}_{err}.$$

The reconstructed quantised transform coefficients l^{12}_{rec} are then supplied to inverse quantisation block 1230 where they are inverse quantised to form inverse quantised reconstructed transform coefficients d^{12}_{rec}. The inverse quantised reconstructed transform coefficients d^{12}_{rec} are then passed to inverse transform block 1240 where they are subjected to an inverse transform operation (e.g. an Inverse Discrete Cosine Transform (IDCT)). As a result, a block of reconstructed pixel values is formed for the current block of the secondary S-frame S_{12}. The reconstructed pixel values $I_c(x,y)$ are supplied to the video output and to frame memory 1250. As subsequent blocks of the secondary S-frame S_{12} are encoded, transmitted from encoder 1100 to decoder 1200 and subsequently decoded, a decoded version of the secondary S-frame is progressively built up in frame memory 1250. From there, the already decoded

blocks of the secondary S-frame can be retrieved and used by intra prediction block 1270 to form predicted pixel values P(x,y) for subsequent blocks of the secondary S-frame S_{12}. Here it should be remembered that the quantised prediction error transform coefficients for each image block of the secondary S-frame S_{12} are generated in encoder 1100 according to the relationship:

$$I^{12}_{err} = I^{2}_{rec} - I^{12}_{pred}$$

where I^{2}_{rec} are quantised reconstructed transform coefficient values produced by encoding and subsequently decoding S-frame S_2 in the second bit-stream. This means that the reconstructed transform coefficients I^{12}_{rec} produced by decoding the compressed video information for secondary S-frame S_{12} are identical to those which would have been produced if S-frame S_2 from the second bit-stream had been transmitted and decoded. As previously explained, this follows since:

$$I^{12}_{rec} = I^{12}_{pred} + I^{12}_{err}$$

$$= I^{12}_{pred} + I^{2}_{rec} - I^{12}_{pred} = I^{2}_{rec}$$

Thus, $I^{12}_{rec} = I^{2}_{rec}$.

Therefore it can be seen that by constructing a secondary S-frame S_{12} according to the method of the invention, transmitting it from encoder to decoder and subsequently decoding it, mis-match-free switching between a first and a second bit-stream can be achieved.

Consider the case that the secondary S-frame is an SI-frame while the S-frame within the bit stream is an SP-frame. In this case, a frame which uses motion-compensated prediction is represented by a frame which uses only spatial prediction. This special case has implications in random access and error resiliency which will described later.

It should be noted that in encoder 1100 and decoder 1200 according to the first embodiment of the invention, described above, the quantisation applied to the transform coefficients c_{pred} produced in transform blocks 1160 (encoder) and 1290 (decoder) to yield quantised transform coefficients l_{pred} is the same as that used to generate the quantised prediction error transform coefficients l_{err}. More specifically, in the first embodiment of the invention, when a block of predicted pixel values P(x,y) is produced for an image block of an S-frame which is being encoded/decoded, the quantisation parameter QP used to quantise the transform coefficients c_{pred} corresponding to the predicted block of pixel values P(x,y) should be the same as the quantisation parameters used in generating the quantised prediction error transform coefficients l_{err}. This is advantageous since the summation performed in order to produce reconstructed transform coefficients l_{rec} is performed in the quantised transform coefficient domain, i.e. since

$$l_{rec} = l_{pred} + l_{err},$$

failure to use identical quantisation parameters in the construction of l_{pred} and l_{err} will lead to errors in the reconstructed quantised transform coefficients l_{rec}.

Figure 15 shows a block diagram of an S-frame encoder 1500 according to a second embodiment of the invention which provides greater flexibility in the choice of quantisation parameters to generate the quantised transform coefficients l_{pred} and l_{err}. As can be seen by comparing Figure 15 with Figure 11, the principal difference between the S-frame encoder 1500 according to a second embodiment of the invention and S-frame encoder 1100 according to a first embodiment of the invention concerns the location of quantisation blocks 1525 and 1550. The operation of the S-frame encoder 1500 according to the second embodiment of the invention will now be described in detail with reference to Figure 15.

According to the second embodiment of the invention a video frame to be encoded in S-frame format is first partitioned into blocks and each

block is then encoded as either on SP-block or on SI-block. Switch 1585 is operated as appropriate to switch between the SP and SI coding modes. In SP coding mode switch 1585 is operated to obtain a motion compensated prediction for the current block of the frame being encoded from motion compensated prediction block 1575. Motion compensated prediction block 1575 forms a block of predicted pixel values P(x,y) for the current block of the frame being encoded by determining a motion vector describing the relationship between the pixels of the current block and pixel values of a reconstructed reference frame held in frame memory 1570.

In SI-coding, mode switch 1585 is operated to obtain a prediction for the current block of the frame being encoded from intra prediction block 1580. Intra prediction block 1580 operates in a manner analogous to that described in connection with the first embodiment of the invention to form a block of predicted pixel values P(x,y) for the current block of the frame being encoded using spatial prediction from already encoded neighbouring pixels within the frame being encoded.

In both SP- and SI-coding modes a forward transform, for example a Discrete Cosine Transform (DCT), is applied to the predicted block of pixel values P(x,y) in transform block 1590. The resulting transform coefficients c_{pred} are supplied to summers 1520 and 1540. The original image data, comprising the actual pixel values of the image block currently being encoded, are passed to transform block 1510 where they too are subjected to a forward transform (e.g DCT). The resulting transform coefficients c_{orig} are then passed to summer 1520 which forms a difference between c_{orig} and c_{pred} to generate prediction error transform coefficients c_{err} according to the relationship:

$$c_{err} = c_{orig} - c_{pred}.$$

The prediction error transform coefficients are supplied to quantisation block 1525 where they are quantised using quantisation parameter PQP to form quantised prediction error transform coefficients l_{err} which are, in turn passed to multiplexer 1540.

If the current block is being encoded in SP-format, multiplexer 1540 also receives information relating to the motion vectors used in formation of the motion compensated prediction P(x,y) for the SP-coded block. If the current block is being encoded in SI-format information concerning the intra prediction mode used to form the prediction P(x,y) for the SI-coded block is also passed to the multiplexer. Advantageously, multiplexer 1540 applies appropriate variable length coding (VLC) to the quantised prediction error transform coefficients l_{err} and to the motion vector or intra prediction mode information and forms a bit-stream for transmission to a corresponding decoder by multiplexing together the various forms of information.

The quantised prediction error transform coefficients l_{err} are passed from quantisation block 1525 to inverse quantisation block 1530 where they are inverse quantised using quantisation parameter PQP to form inverse quantised prediction error transform coefficients d_{err}. The inverse quantised prediction error transform coefficients d_{err} are next passed to summer 1540 where they are combined with the transform coefficients c_{pred} produced from the predicted pixel values P(x,y) for the current block. More specifically, summer 1540 adds transform coefficients c_{pred} and inverse quantised prediction error transform coefficients d_{err} to form reconstructed transform coefficients c_{rec} according to the relationship:

$$c_{rec} = c_{pred} + d_{err}$$

The reconstructed transform coefficients c_{rec} are then passed to quantisation block 1550 where they are quantised using quantisation parameter SPQP to produce quantised reconstructed transform coefficients l_{rec}. It should be noted that quantisation parameter SPQP used to quantise the reconstructed transform coefficients is not necessarily the same as the quantisation parameter PQP used to quantise the prediction error transform coefficients c_{err} in quantisation block 1525. In particular, a finer quantisation can be applied to reconstructed transform coefficients c_{rec} and a courser quantisation can

be applied to prediction error coefficients c_{err}. This ultimately results in a smaller reconstruction error (distortion) when a decoded image is formed in the decoder.

Quantised reconstructed transform coefficients l_{rec} are next supplied to inverse quantisation block 1560 where they are inverse quantised using quantisation parameter SPQP to form inverse quantised reconstructed transform coefficients d_{rec}. The inverse quantised reconstructed transform coefficients d_{rec} are then passed to inverse transform block 1565 where they are subjected to an inverse transform operation e.g. an Inverse Discrete Cosine Transform (IDCT). As a result of this operation, a block of reconstructed pixel values $l_c(x,y)$ is formed for the image block in question. The block of reconstructed pixel values $l_c(x,y)$ is subsequently stored in frame memory 1570. As subsequent blocks of the frame being encoded in S-frame format undergo the previously described encoding and local decoding operations, a decoded version of the current frame is progressively assembled in frame memory 1570 from where it can be accessed and used in intra prediction of subsequent blocks of the same frame or in inter (motion compensated) prediction of subsequent frames in the video sequence.

Operation of an S-frame decoder 1300 according to the second embodiment of the invention will now be described with reference to Figure 13. The bit-stream generated by the S-frame encoder 1500 according to the second embodiment of the invention and described above in connection with Figure 15 is received by decoder 1300 and is demultiplexed into its constituent parts. The decoder reconstructs a decoded version of the S-frame on a block-by-block basis. As previously described, an S-frame generally comprises both SP-coded and SI-coded image blocks. For SP-coded image blocks, the information in the received bit-stream comprises VLC encoded motion vector information and VLC encoded quantised prediction error transform coefficients l_{err}. For image blocks encoded in SI-format, the information in the received bit-stream comprises VLC encoded information relating to the intra prediction mode used to form the intra

prediction for the SI-coded block, as well as VLC encoded quantised prediction error transform coefficients I_{err}.

When decoding an SP-coded image block, demultiplexer 1310 first applies appropriate variable length decoding (VLD) to the received bit-stream to recover the motion vector information and quantised prediction error coefficients I_{err}. It then separates the motion vector information from the quantised prediction error coefficients I_{err}. The motion vector information is supplied to motion compensated prediction block 1370 and the quantised prediction error coefficients I_{err} recovered from the received bit-stream are applied to inverse quantisation block 1320. The motion vector information recovered from the received bit-stream is used in motion compensated prediction block 1370 together with pixel values of a previously reconstructed frame held in frame memory 1360 to form a prediction P(x,y) for the current block being decoded, in a manner analogous to that employed in encoder 1500.

When decoding an SI-coded image block, demultiplexer 1310 applies appropriate variable length decoding to the received intra prediction mode information and the quantised prediction error transform coefficients I_{err}. The intra prediction mode information is then separated from the quantised prediction error transform coefficients I_{err} and is supplied to intra prediction block 1380. The quantised prediction error transform coefficients I_{err} are supplied to inverse quantisation block 1320. The intra prediction mode information recovered from the received bit-stream is used in intra prediction block 1380 in conjunction with previously decoded pixel values of the current frame held in frame memory 1360 to form a prediction P(x,y) for the current block being decoded. Again, the intra prediction process performed in decoder 1200 is analogous to that performed in the corresponding encoder 1500 previously described.

For both SP- and SI- coded image blocks the quantised prediction error transform coefficients I_{err} recovered from the received bit-stream are inverse quantised in inverse quantisation block 1320 using quantisation parameter PQP to form inverse quantised prediction error transform

coefficients d_{err}. The inverse quantised prediction error transform coefficients d_{err} are applied to one input of summer 1325.

Once a prediction P(x,y) for the current block of the frame being decoded has been formed, either by motion compensated prediction in motion compensated prediction block 1370 or by intra prediction in intra prediction block 1380, switch 1385 is operated as appropriate to supply the predicted pixel values P(x,y) to transform block 1390. Here a forward transform, for example a Discrete Cosine Transform (DCT) is applied to the predicted block of pixel values P(x,y) to form transform coefficients c_{pred}. Transform coefficients c_{pred} are then supplied to the second input of summer 1325 where they are combined with the inverse quantised prediction error transform coefficients received from inverse quantisation block 1320, to form reconstructed transform coefficients c_{rec}. More specifically, the reconstructed transform coefficients are determined by adding together transform coefficients c_{pred} and inverse quantised prediction error transform coefficients d_{err} according to the relationship:

$$c_{rec} = c_{pred} + d_{err}.$$

The reconstructed transform coefficients c_{rec} are then passed to quantisation block 1330 where they are quantised to produce quantised reconstructed transform coefficients l_{rec} using quantisation parameter SPQP. The quantised reconstructed transform coefficients l_{rec} are, in turn, supplied to inverse quantisation block 1340 where they are inverse quantised using quantisation parameter SPQP to form inverse quantised reconstructed transform coefficients d_{rec}. The inverse quantised reconstructed transform coefficients d_{rec} are then passed to inverse transform block 1350 where they are subjected to an inverse transform operation, for example an Inverse Discrete Cosine Transform (IDCT). As a result of the inverse transformation applied in inverse transform block 1350, a block of reconstructed image pixels $l_c(x,y)$ is formed for the image block in question. The block of reconstructed pixels $l_c(x,y)$ is supplied to the video output of the decoder and to frame memory 1360, where the pixels are stored. As subsequent blocks of

the S-frame undergo the decoding operations described above, a decoded version of the current frame is progressively assembled in frame memory 1360. From there in can be accessed and used in the intra prediction of subsequent blocks of the same frame or in inter (motion compensated) prediction of subsequent frames in the video sequence.

In Fig. 16, an encoder according to a third embodiment of the invention is shown. In this embodiment, the transform coefficients c_{pred} undergo quantisation and inverse quantisation using the same quantisation parameter SPQP both in the encoder section (blocks 1625 and 1630) and in the decoder section (blocks 1692 and 1694). Therefore, the encoder does not introduce any additional quantisation error to the prediction loop, and error build-up in the prediction loop is thus effectively prevented. The blocks 1610, 1620, 1625, 1630, 1640, 1650, 1660, 1665, 1670, 1675, 1680, 1685, 1690 have similar functionality to the blocks 1510, 1520, 1525, 1530, 1540, 1550, 1560, 1565, 1570, 1575, 1580, 1585, 1590 presented in Fig. 15, respectively.

In Fig. 6, a decoder 600 according to an advantageous embodiment of the invention is described. Most of the elements of the decoder 600 are the same than of the decoder 1200 presented in Fig. 12. The operational blocks of the decoder 600 are arranged to decode the prediction blocks of the frames, wherein no switching means are not shown in Fig. 6. Other blocks 610, 615, 620, 630, 640, 650, 660 and 670 have similar functionality to the blocks 1210, 1220, 1230, 1240, 1250, 1260, 1290 and 1295 presented in Fig. 12, respectively.

In Fig. 9, a decoder 600 according to another advantageous embodiment of the invention is described. The decoder 600 presented in Fig. 9 is a modification from the decoder 600 shown in Fig. 6. The difference between the decoder presented in Fig. 9 and the decoder presented in Fig. 12 is that a normalisation block 680 is inserted between the demultiplexer 610 and one input of summing element 615. Other blocks 610, 615, 620, 630, 640, 650, 660 and 670 have similar

functionality to the blocks 1210, 1220, 1230, 1240, 1250, 1260, 1290 and 1295 presented in Fig. 12, respectively.

In Fig. 10, a decoder 600 according to yet another advantageous embodiment of the invention is described. Most of the elements of the decoder 600 are the same than of the decoder 1300 presented in Fig. 13. The operational blocks of the decoder 600 are arranged to decode the prediction blocks of the frames, wherein no switching means are not shown in Fig. 10. The other difference between the decoder presented in Fig. 13 and the decoder presented in Fig. 10 is that a normalisation block 680 is used instead of inverse quantisation block 1230. Other blocks 610, 615, 620, 630, 640, 650, 660 and 670 have similar functionality to the blocks 1310, 1325, 1330, 1340, 1350, 1360, 1370 and 1390 presented in Fig. 13, respectively.

The encoding of a video frame can be performed block-by-block basis such that in the same encoded video frame there can exist differently encoded areas. For example, some parts of the frame can be inter coded and some other parts of the frame can be intra coded. The above procedures are applied to each part of the frame appropriately, according to the encoding procedure of the part in question.

In addition to the transmission network, the request for the change of the bit stream transmission properties may also be originated by other parts of the transmission system. For example, the receiver may request the transmitting server to change the parameters for some reason. This request is delivered to the transmitting server *e.g.* via the transmission network.

Although H.26L is used as an example of a standard, embodiments of the present invention and any variations and modifications are deemed to be within the scope of the present invention.

Bit stream switching is not the only application in which the present invention can be applied. If one of the bit streams has a lower temporal resolution, e.g. 1 frame/sec, this bit stream can be used to provide fast-

forward functionality. Specifically, decoding from the bit stream with a lower temporal resolution and then switching to the bit stream with a normal frame rate would provide such functionality. Fig. 8 depicts two bit streams the second of which comprises only S-frames predicted from each other at intervals greater than the frame repetition interval of the first bit-stream. Furthermore, "Fast Forward" can start and stop at any location in the bit-stream. In the following, some other applications of the present invention are described.

Splicing and Random Access

The bit stream-switching example discussed earlier considered bit streams belonging to the same sequence of images. However, this is not necessarily the case in all situations where bit stream switching is needed. Examples include: switching between bit streams arriving from different cameras capturing the same event but from different perspectives, or cameras placed around a building for surveillance; switching to local/national programming or insertion of commercials in a television broadcast, video bridging, etc. The general term for the process of concatenating encoded bit streams is splicing.

When switching occurs between bit streams that belong to different sequence of images, this only affects the encoding of the S-frames used to switch between bit streams, i.e., the secondary S-frame S_{12} in Fig. 5. Specifically, the use of motion-compensated prediction of frames in one sequence of images using reference frames from a different sequence of images is not as efficient as when both bit streams belong to the same sequence of images. In this case, it is most likely that spatial prediction of the secondary S-frames frames would be more efficient. This is illustrated in Fig. 7 where the switching frame is an SI-frame using only spatial prediction, which identically reconstructs the corresponding SP-frame S_2. This method can be used as a random access mechanism into a bit stream and has further implications in error recovery and resilience as described in the following.

Error recovery

Multiple representations of a single frame in the form of S-frames predicted from different reference frames, e.g., predicted from the immediately preceding reconstructed frame and a reconstructed frame further back in time can be used to increase error resilience of an encoded video sequence and/or to improve recovery from errors in the bit-stream. This is illustrated in Fig. 14. In a situation where a packet loss occurs during streaming of a pre-encoded bit stream and a frame or slice is lost, the receiver informs the transmitter of the lost frame/slice and the transmitter responds by sending one of the alternative representations of the next S-frame. The alternative representation, *e.g.* frame S_{12} in Fig. 14, uses reference frames that have already been correctly received by the receiver. For slice based packetization and delivery, the sender could further estimate the slices that would be affected by such a slice/frame loss and update only those slices in the next S-frame with their alternative representations.

Similarly, and as mentioned earlier in the discussion of splicing, the secondary representation of the S-frame can be generated without using any reference frames, i.e., an SI_2-frame such as that shown in Fig. 14. In this case, the sender would send the secondary SI-frame, i.e., SI_2 instead of S_2 to stop error propagation. This approach could also be extended in a straight forward manner to slice-based encoding/packetization. More specifically, the server sends the slices in the next S-frame, which would be affected by the packet loss, from SI-frame.

Error resilience

The encoding of a video frame can be performed block-by-block basis such that in the same encoded video frame there can exist differently encoded areas. For example, some parts of the frame can be inter coded and some other parts of the frame can be intra coded. And as already discussed, the intra-block coding, since it doesn't employ any

temporal correlation, stops any error propagation which might be started due to transport impairments.

In lossy transport networks, an intra macroblock refresh strategy can provide significant error resiliency/recovery performance. In an interactive client/server scenario, the encoder on the server side decides to encode the frames/macroblocks either based on specific feedback received from the client, e.g., the exact location of the lost/corrupted frame/slice/macro block, or based on the expected network conditions calculated through negotiation or the measured network conditions. This kind of intra-macroblock update strategy improves the quality of the received video by providing error resiliency and error recovery. The optimal intra-macroblock update refresh rate, i.e., the frequency at which macroblocks are intra-coded, depends on the transport channel conditions, e.g., packet loss and/or bit error rate. However, when already encoded bit streams are sent, which is the case in typical streaming applications, the above strategy cannot be applied directly. Either the sequence needs to be encoded with the worst-case expected network conditions or additional error resiliency/recovery mechanisms are required.

From the earlier discussion concerning the use of S-frames in error recovery and splicing applications, it can be noted that S-frames or slices within them can easily be represented as SI-frames/slices that do not use any reference frames while still leading to identical reconstruction of the S-frame. This feature can be exploited in the adaptive intra refresh mechanism discussed above. First, a sequence of images is encoded with some predefined ratio of S-macroblocks. Then during the transmission, some of the S-macroblocks are sent in their secondary representation that is as SI-macro blocks. The number of S-macroblocks to be sent in SI representation can be calculated in a manner similar to the method used in the real-time encoding/delivery approach previously described.

Video Redundancy Coding

S-frames have other uses in applications in which they do not act as replacements of I-frames. Video Redundancy Coding (VRC) can be given as an example. The principle of the VRC method is to divide a sequence of pictures into two or more threads in such a way that all pictures in the sequence are assigned to one of the threads in a round-robin fashion. Each thread is coded independently. At regular intervals, all threads converge into a so-called sync frame. From this sync frame, a new thread series is started. If one of the threads is damaged e.g. because of a packet loss, the remaining threads typically stay intact and can be used to predict the next sync frame. It is possible to continue the decoding of the damaged thread, which leads to slight degradation in picture quality, or to stop the decoding of the damaged thread, which leads to a drop in the frame rate. Sync frames are always predicted from one of the undamaged threads. This means that the number of transmitted I-frames can be kept small, because there is no need for complete re-synchronization. For the sync frame more than one representation (P-frame) is sent, each one using a reference frame from a different thread. Due to the use of P-frames these representations are not identical. Therefore mismatch is introduced when some of the representations cannot be decoded and their counterparts are used when decoding the following threads. The use of S-frames as sync frames eliminates this problem.

It is obvious that the present invention is not limited to the above described embodiments but it can be modified within the scope of the appended claims.

Claims:

1. A method for transmitting video information, in which at least a first bit-stream (510) and a second bit-stream are formed from the video information, the first bit-stream (510) comprises a first set of frames comprising at least one video frame, and the second bit-stream (520) comprises a second set (520) of frames comprising at least one predictive video frame (524), at least partly different encoding parameters are used with encoding of the frames of said first bit-stream (510) and said second bit-stream (520), at least one frame of said first bit-stream (510) is being transmitted, and the transmission is switched over from said first bit-stream (510) to said second bit-stream (520), **characterized** in that in switching the transmission over from said first bit-stream (510) to said second bit-stream (520), the secondary bit-stream comprises at least one primary switching frame (523), a secondary switching frame (550) is transmitted, the secondary switching frame (550) having been encoded using the encoding parameters of the second bit-stream (520) and at least one reference frame from the first bit-stream (510), and said secondary switching frame (550) is used in place of the primary switching frame (523) as a reference frame to be used in the reconstruction of said at least one predictive video frame (524) of the second set of video frames.

2. The method according to claim 1, **characterized** in that the first bit-stream (510) comprises at least one primary switching frame (513).

3. The method according to claim 1 or 2, **characterized** in that the first bit-stream (510) comprises only one intra frame and one primary switching frame (513) to perform a changeover from one location of the video information to another location of the video information.

4. The method according to claim 1 or 2, **characterized** in that the first bit-stream (510) comprises only intra frames and primary switching frames (513) to perform fast forward operation to the video information.

5. The method according to any of the claims 2 to 4, **characterized** in that said primary switching frame (523) is a predictive video frame (524), in which prediction information comprises only intra prediction information.

6. The method according to any of the claims 1 to 5, **characterized** in that the primary switching frame (513) is formed so that transform coefficients c_{pred} are calculated and quantised to form quantised values of the transform coefficients l_{pred}, that quantised coefficients l_{err} for prediction error are defined, that dequantised reconstruction image coefficients d_{rec} are defined so that coefficients c_{rec} exist from which by quantisation and dequantisation, dequantised reconstruction image coefficients d_{rec} can be obtained, and that dequantised reconstruction image coefficients d_{rec} fulfil one of the following conditions:

$d_{rec} = d_{pred} + d_{err}$, or
$c_{rec} = c_{pred} + d_{err}$,

in which d_{err} are dequantised values of the prediction error, and
d_{pred} are the dequantised values of the quantised transform coefficients l_{pred}.

7. The method according to claim 6, **characterized** in that same quantisation parameters are used with quantisation of the transform coefficients l_{pred}, and with quantisation of prediction error.

8. The method according to claim 6, **characterized** in that different quantisation parameters are used with quantisation of the transform coefficients l_{pred}, than with quantisation of prediction error.

9. The method according to any of the claims 1 to 8, **characterized** in that said secondary switching frame (550) is used for recovering from transmission errors, wherein said primary switching frame (523) is a predictive video frame (524), in which prediction information comprises prediction information from previous video frames than the preceding frame of the predictive video frame (524).

10. The method according to any of the claims 1 to 8, **characterized** in that said secondary switching frame (550) is used for recovering from transmission errors, wherein said primary switching frame (523) is a predictive video frame (524), in which prediction information comprises only intra prediction information.

11. The method according to any of the claims 1 to 10, **characterized** in that both said at least one primary switching frame (523) and said secondary switching frame (550) produce the same reconstruction result of said at least one predictive video frame (524).

12. An encoder comprising means for forming at least a first bit-stream (510) and a second bit-stream from a video information, the first bit-stream comprises a first set of frames comprising at lest one video frame, and the second bit-stream comprises a second set (520) of frames comprising at least one predictive video frame (524), and means for using at least partly different encoding parameters with encoding of the frames of said first bit-stream (510) and said second bit-stream (520), means for transmitting at least one frame of said first bit-stream (510), and means for switching the transmission over from said first bit-stream (510) to said second bit-stream (520), **characterized** in that said means for switching the transmission over from said first bit-stream (510) to said second bit-stream (520) comprise means for encoding a secondary switching frame (550) using the encoding parameters of the second bit-stream (520) and reference frames from the first bit-stream (510) to enable to switch the transmission over from said first bit-stream (510) to said second bit-stream (510).

13. The encoder according to claim 12, **characterized** in that it comprises means (1670, 1675) for producing prediction information by using said reference frames, and means (1692, 1694) for performing quantisation and dequantisation to the prediction information.

14. A decoder for decoding a video information from a signal comprising frames from at least a first bit-stream (510) and a second bit-stream, which are formed from the video information, the first bit-stream comprises a first set of frames comprising at lest one video frame, and the second bit-stream comprises a second set (520) of frames comprising at least one predictive video frame (524), and at least partly different encoding parameters are used with encoding of the frames of said first bit-stream (510) and said second bit-stream (520), **characterized** in that the decoder comprises means for decoding a secondary switching frame (550), said secondary switching frame (550) having been encoded using the encoding parameters of the second bit-stream (520) and at least one reference frame from the first bit-stream (510), and added to said signal in place of a primary switching frame (523) as a reference frame to be used in the reconstruction of said at least one predictive video frame (524) of the second set of video frames, and that said means for decoding a secondary switching frame (550) comprise means for using decoding parameters of the second bit-stream (520) and reference frames from the first bit-stream (510).

15. The decoder according to claim 14, **characterized** in that the primary switching frame (513) comprises areas encoded by intra prediction using only spatial correlation, and areas encoded by inter prediction using motion compensation, wherein the decoder comprises means for using motion compensation information with reconstruction, means for using spatial correlation information with reconstruction, and switching means to perform the reconstruction of each area either by said means for using motion compensation information or by said means for using spatial correlation information, depending on the prediction method used by encoding of the respective area.

16. A signal representing encoded video information, and comprising frames from at least a first bit-stream (510) and a second bit-stream, which are formed from the video information, the first bit-stream comprises a first set of frames comprising at lest one video frame, and the second bit-stream comprises a second set (520) of frames

comprising at least one predictive video frame (524), at least partly different encoding parameters are used with encoding of the frames of said first bit-stream (510) and said second bit-stream (520), **characterized** in that in switching the transmission over from said first bit-stream (510) to said second bit-stream (520), the secondary bit-stream comprises at least one primary switching frame (523), said signal comprises a secondary switching frame (513, 523) having been encoded using the encoding parameters of the second bit-stream (520) and at least one reference frame from the first bit-stream (510), and said secondary switching frame (550) is in place of the primary switching frame (523) as a reference frame to be used in the reconstruction of said at least one predictive video frame (524) of the second set of video frames.

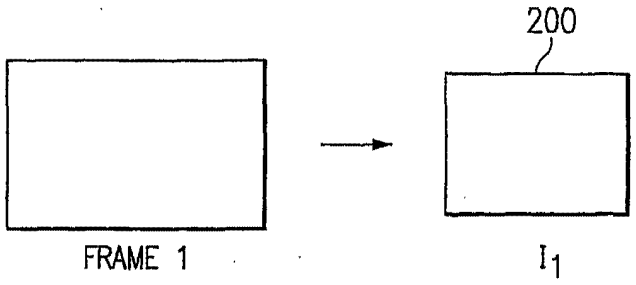


FIG. 1A
(PRIOR ART)

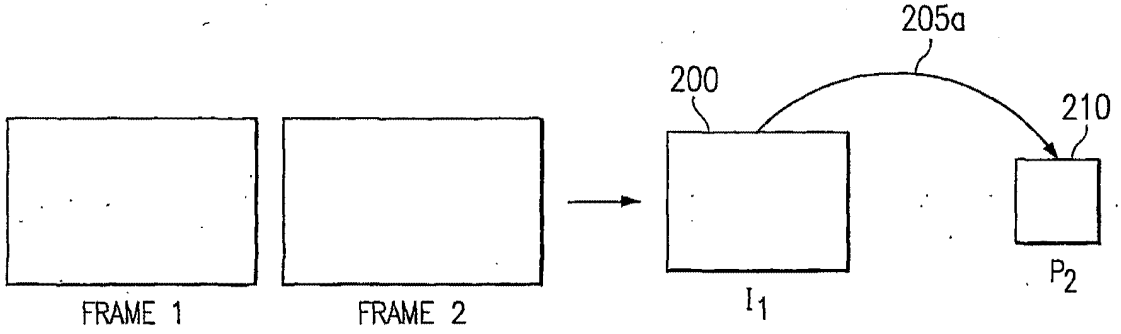


FIG. 1B
(PRIOR ART)

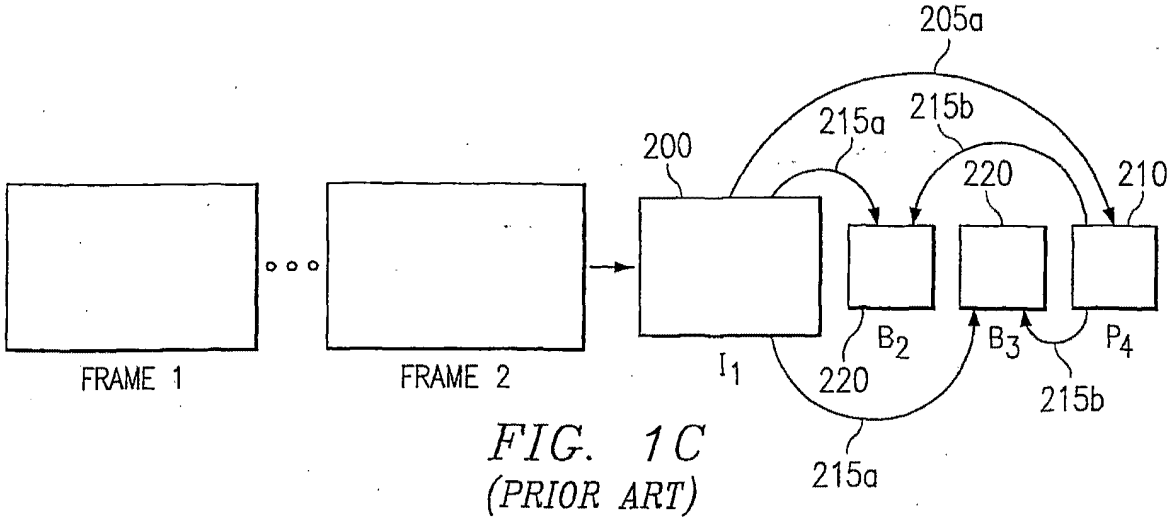


FIG. 1C
(PRIOR ART)

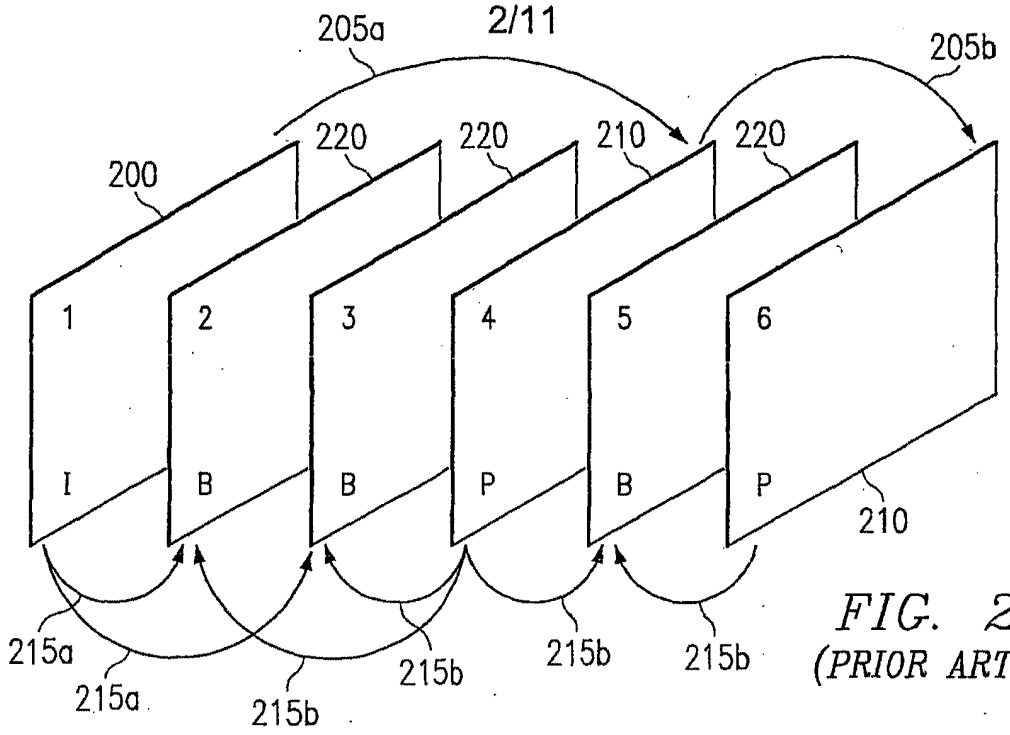


FIG. 2
(PRIOR ART)

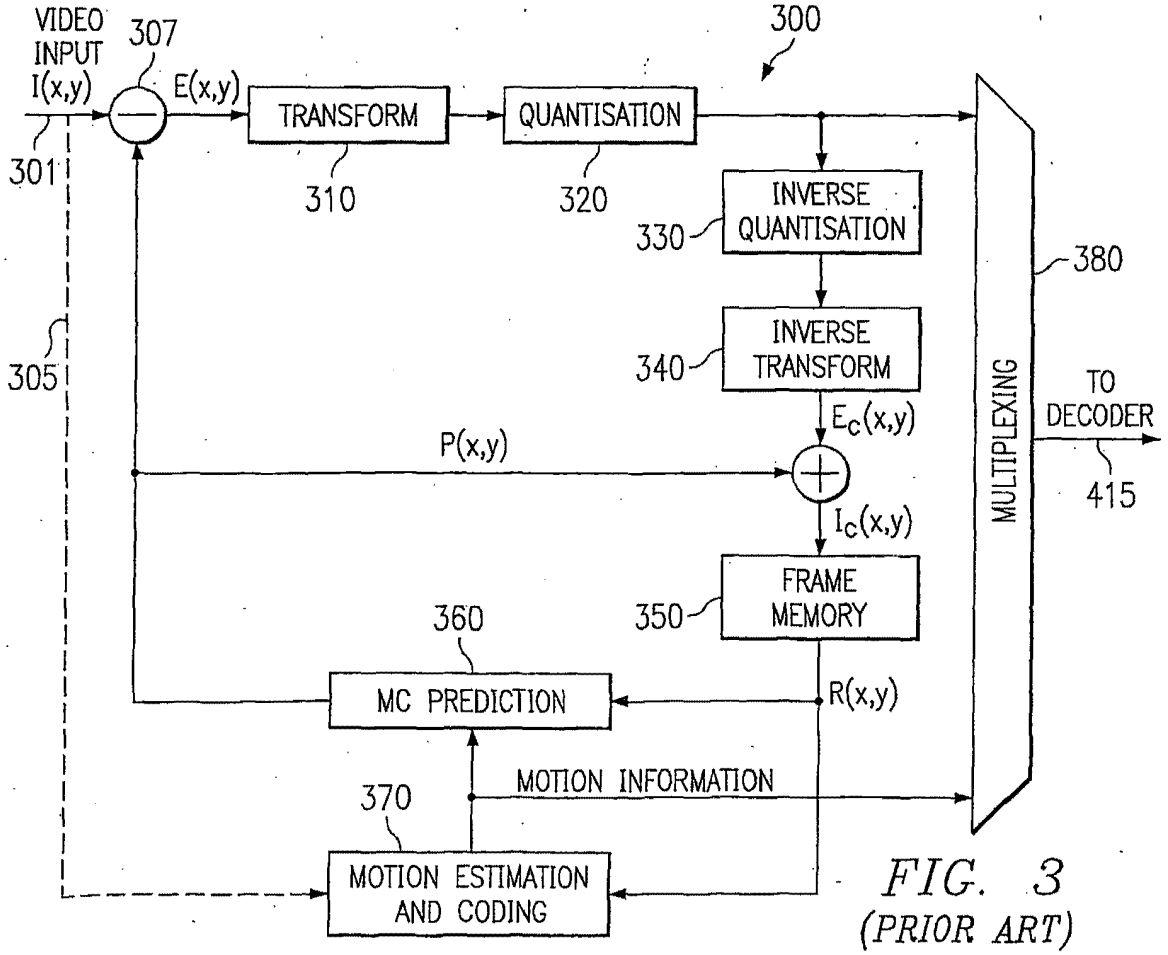


FIG. 3
(PRIOR ART)

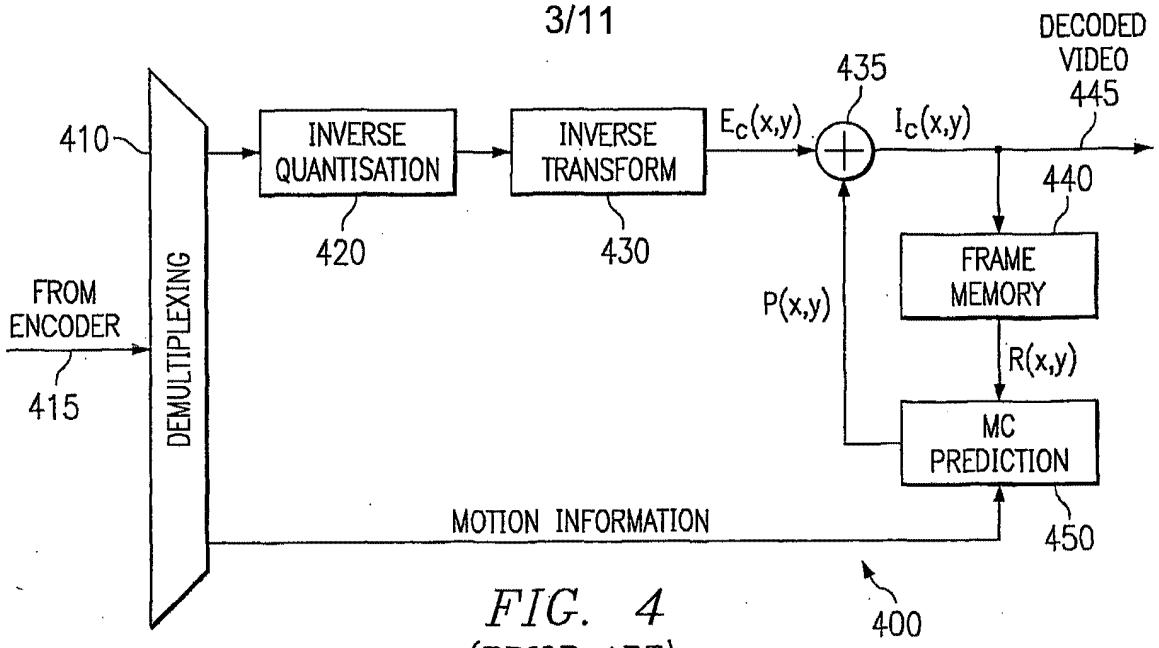


FIG. 4
(PRIOR ART)

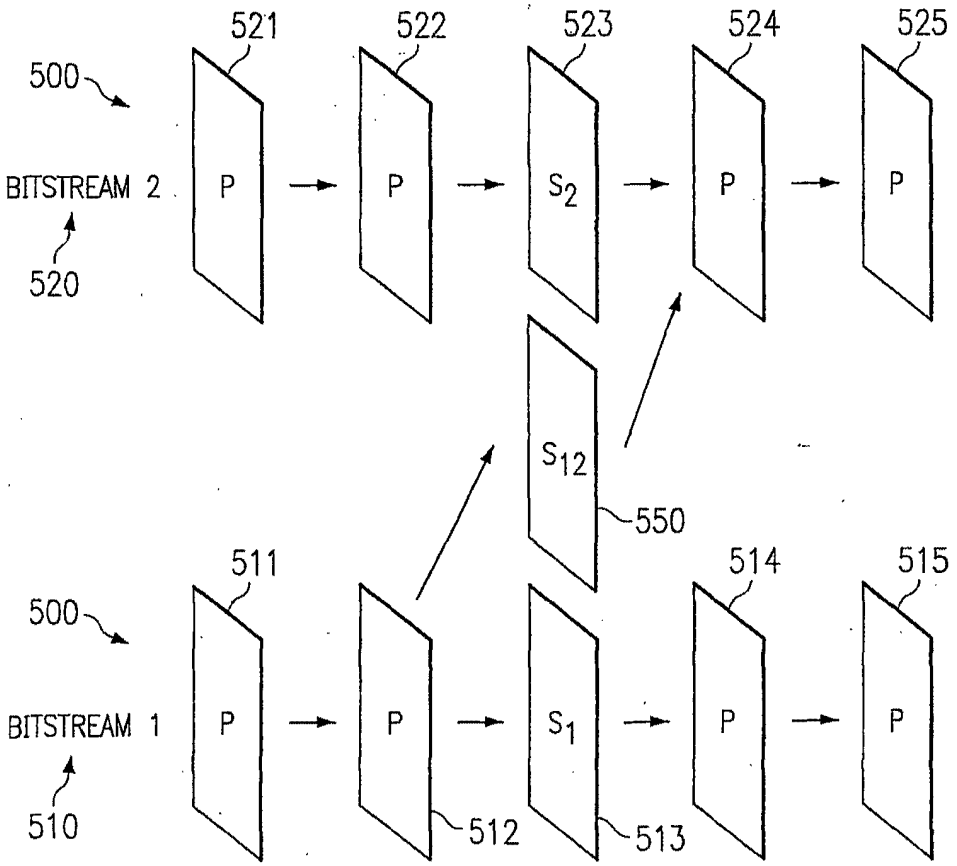


FIG. 5

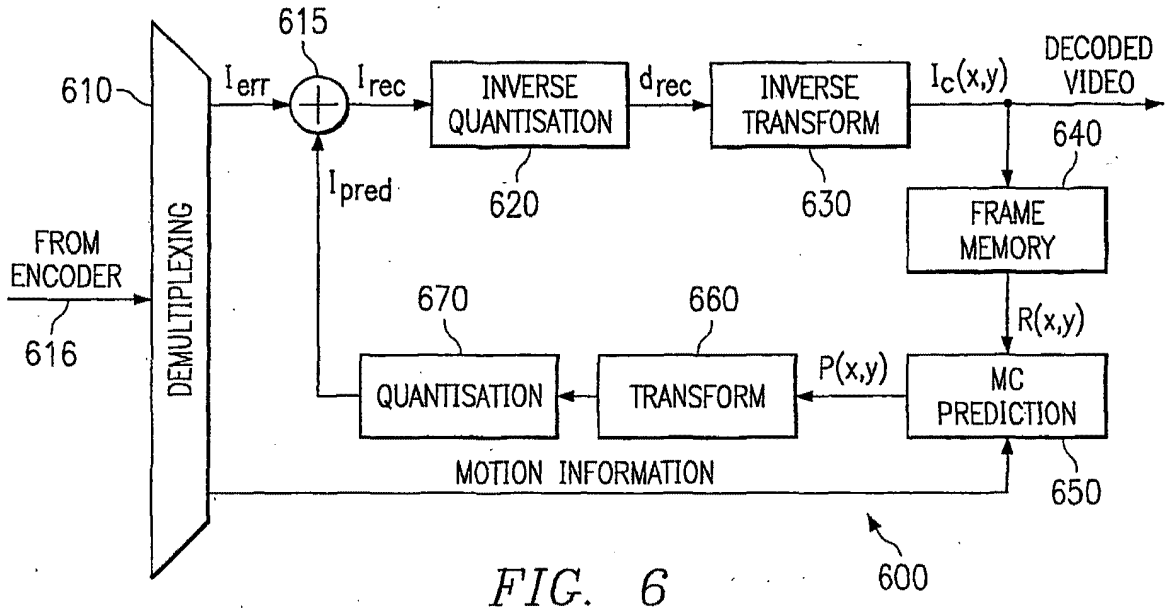

615
610
I_{err}
I_{rec}
INVERSE QUANTISATION
d_{rec}
INVERSE TRANSFORM
$I_C(x,y)$
DECODED VIDEO
640
I_{pred}
620
630
FRAME MEMORY
FROM ENCODER
DEMULTIPLEXING
670
660
R(x,y)
616
QUANTISATION
TRANSFORM
P(x,y)
MC PREDICTION
MOTION INFORMATION
650
600

FIG. 6

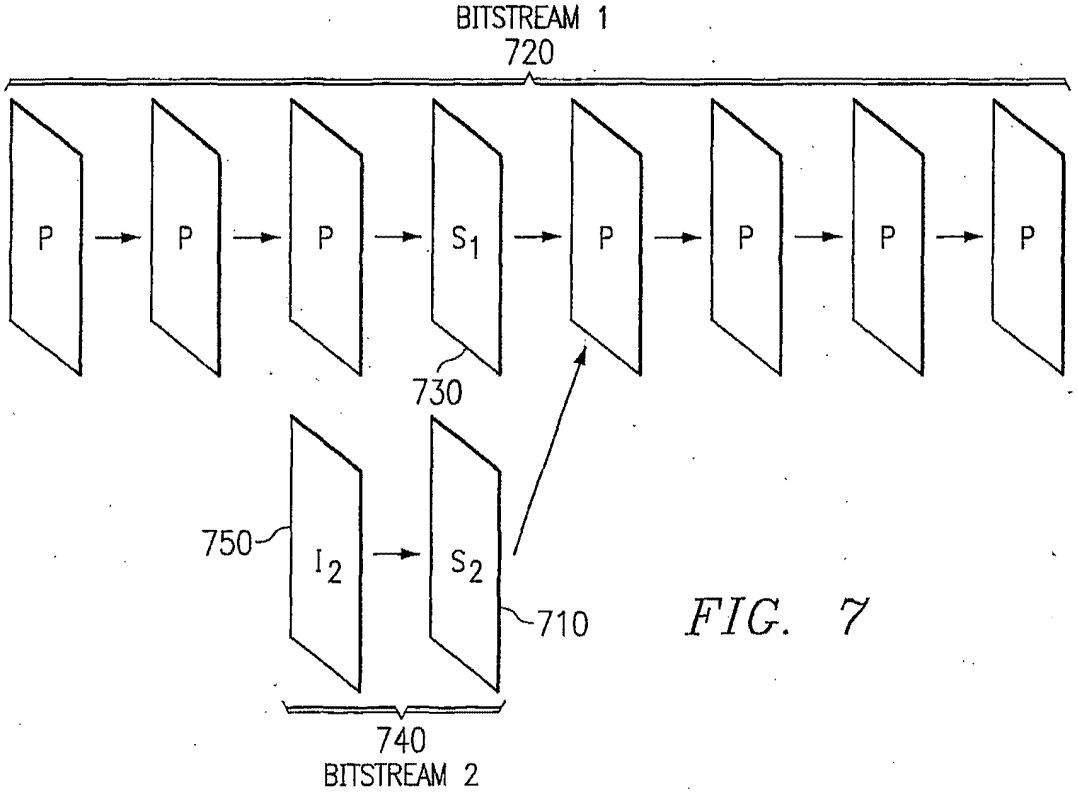

BITSTREAM 1
720
P
P
P
S_1
P
P
P
P
730
750
I_2
S_2
710
740
BITSTREAM 2

FIG. 7

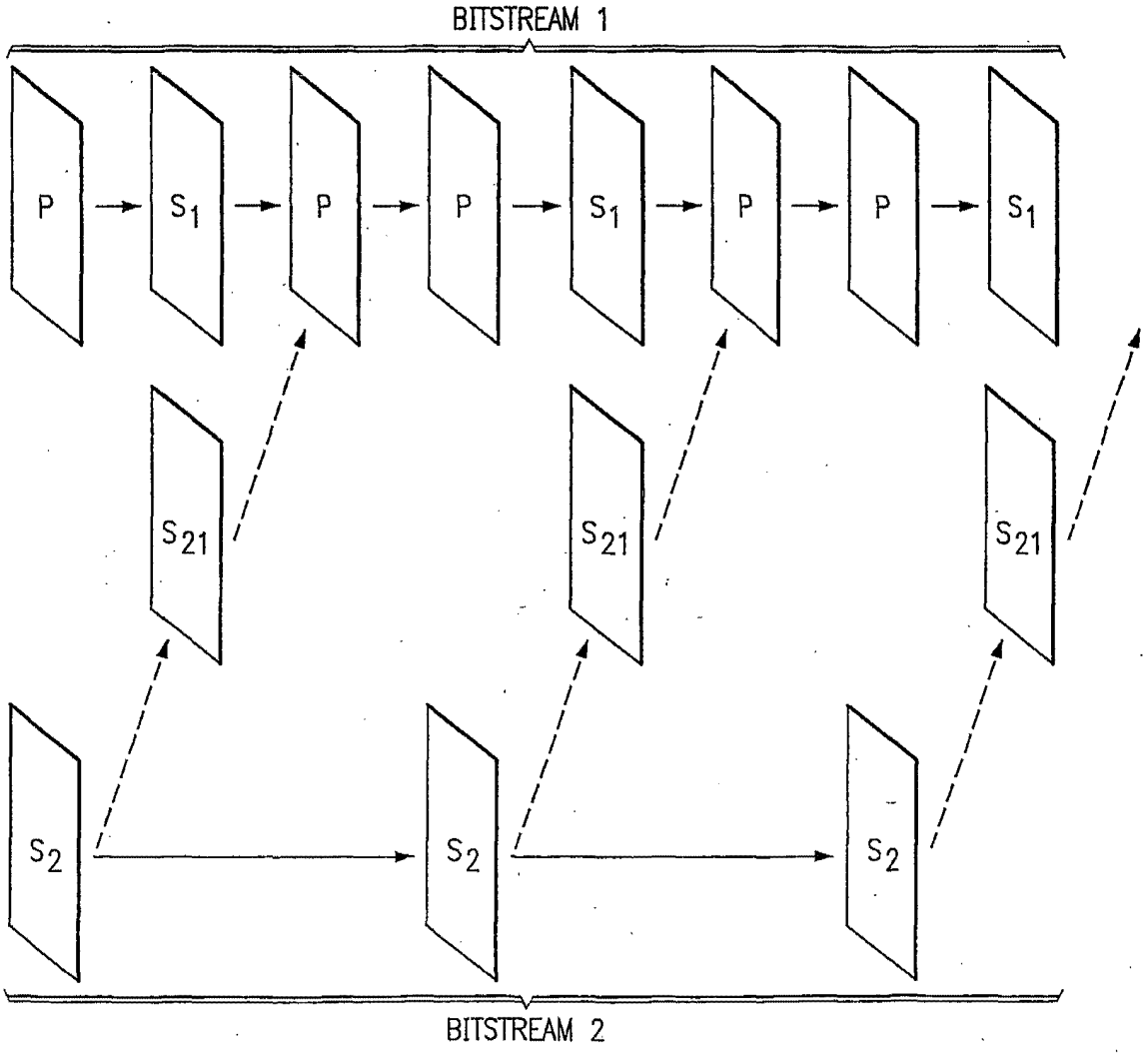


FIG. 8

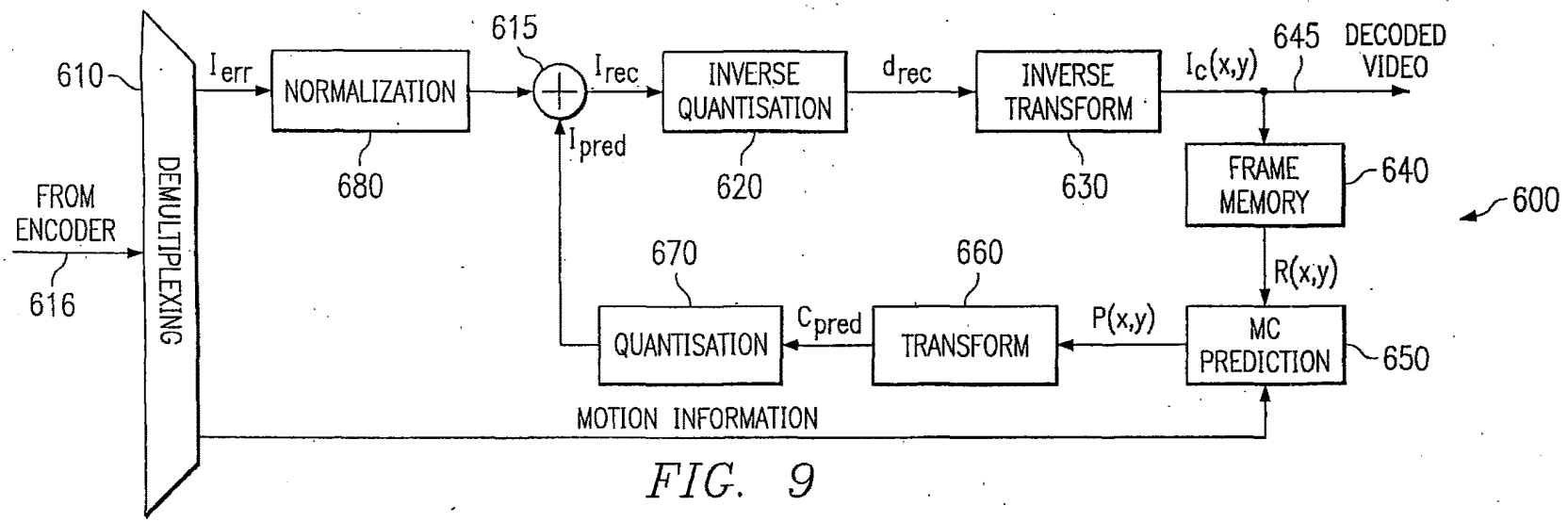


FIG. 9

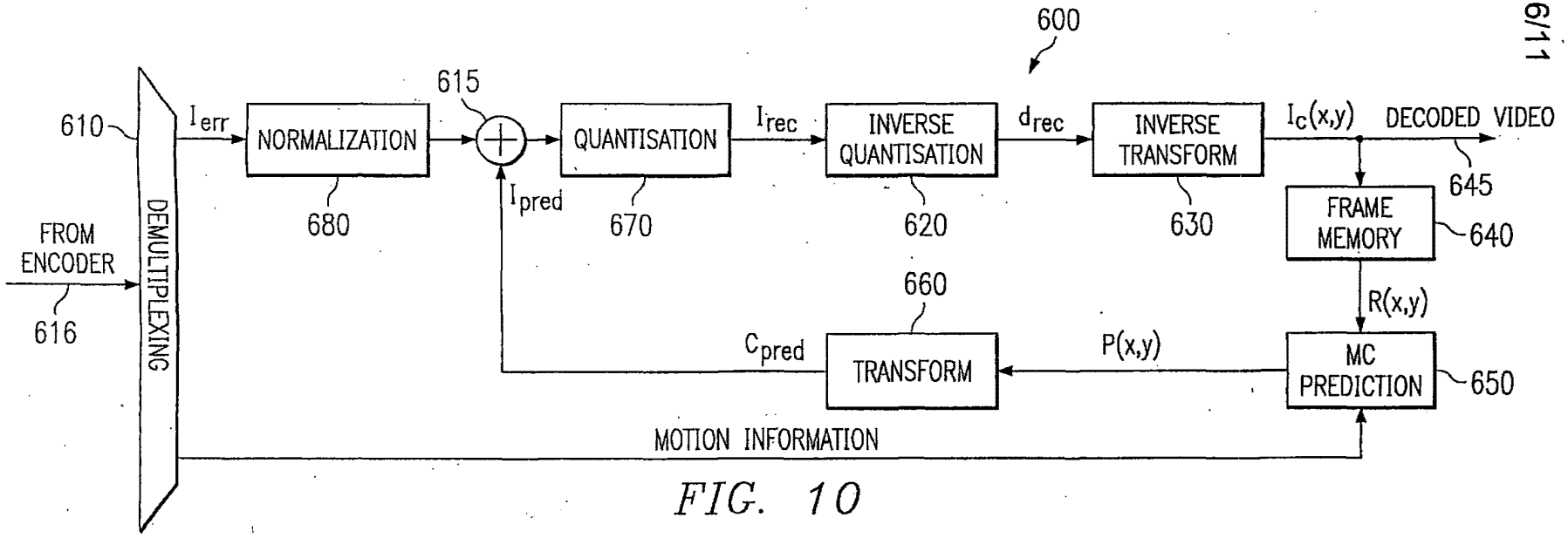


FIG. 10

6/11

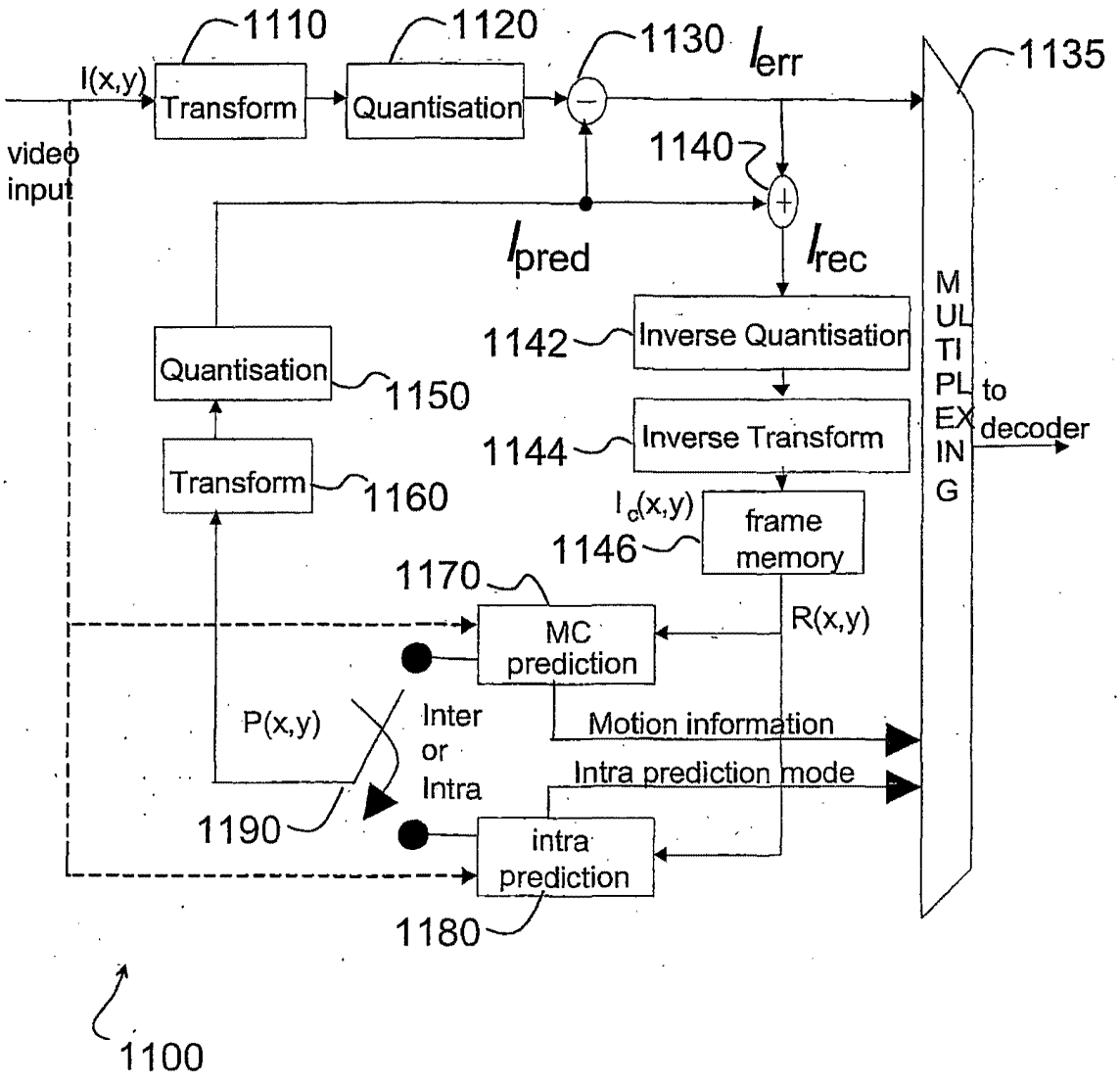

1110
1120
1130
I_{err}
1135
I(x,y)
Transform
Quantisation
video input
1140
I_{pred}
I_{rec}
MULTIPLEXING
to decoder
1142
Inverse Quantisation
Quantisation
1150
1144
Inverse Transform
Transform
1160
$I_{\alpha}(x,y)$
frame memory
1146
1170
R(x,y)
MC prediction
P(x,y)
Inter or Intra
Motion information
Intra prediction mode
1190
intra prediction
1180
1100

FIG. 11

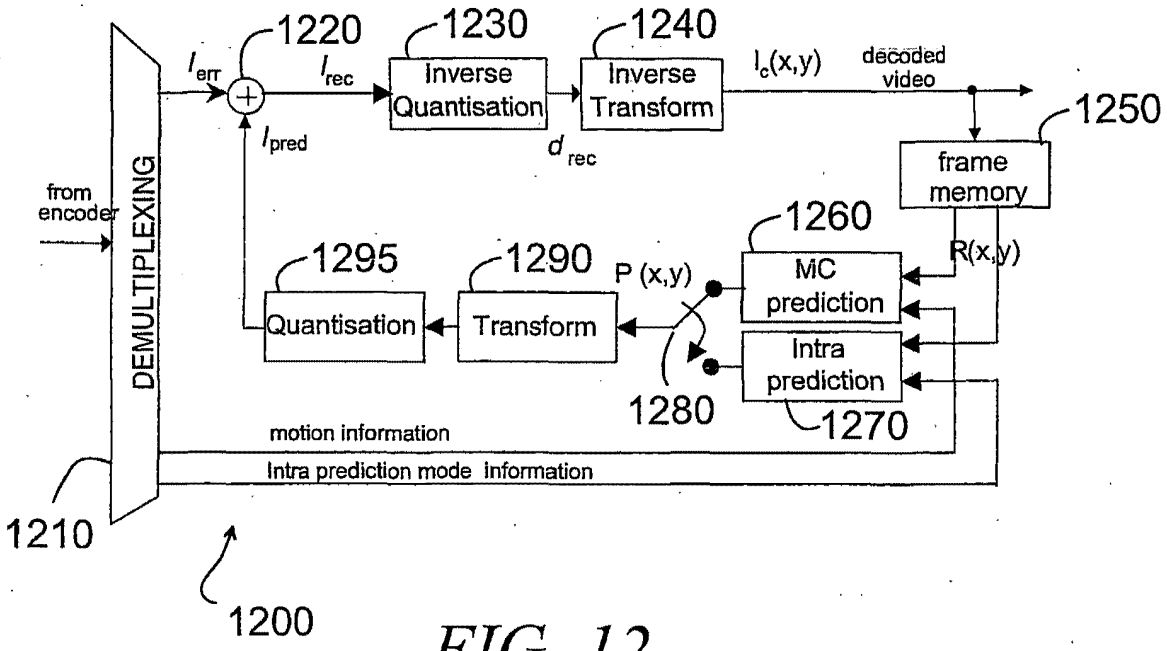

1220
1230
1240
I_{err}
I_{rec}
Inverse Quantisation
Inverse Transform
$I_c(x,y)$
decoded video
1250
I_{pred}
d_{rec}
frame memory
from encoder
DEMULTIPLEXING
1260
1295
1290
P (x,y)
R(x,y)
MC prediction
Quantisation
Transform
Intra prediction
1280
1270
motion information
Intra prediction mode information
1210
1200

FIG. 12

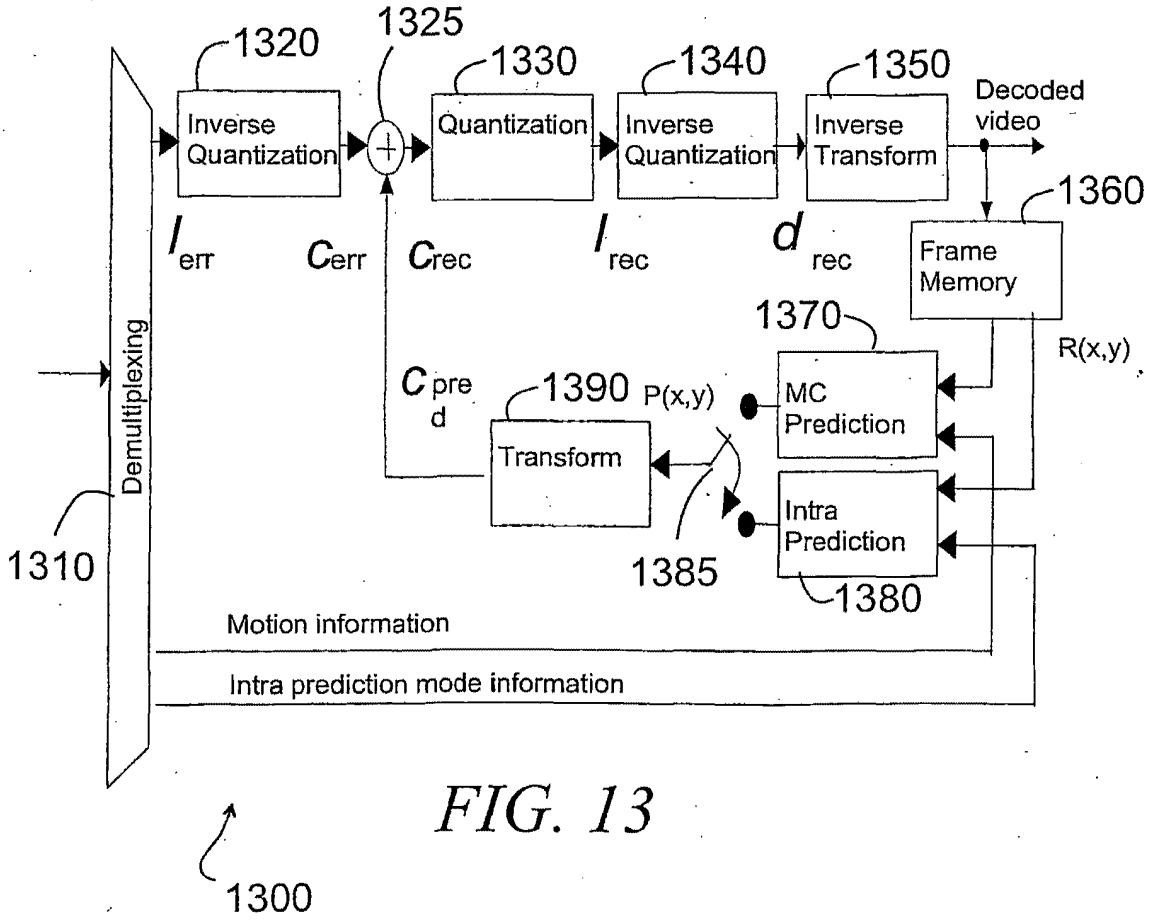

1320
1325
1330
1340
1350
Inverse Quantization
Quantization
Inverse Quantization
Inverse Transform
Decoded video
1360
I_{err}
C_{err}
C_{rec}
I_{rec}
d_{rec}
Frame Memory
1370
R(x,y)
C_{pred}
1390
P(x,y)
MC Prediction
Transform
Intra Prediction
Demultiplexing
1385
1380
1310
Motion information
Intra prediction mode information
1300

FIG. 13

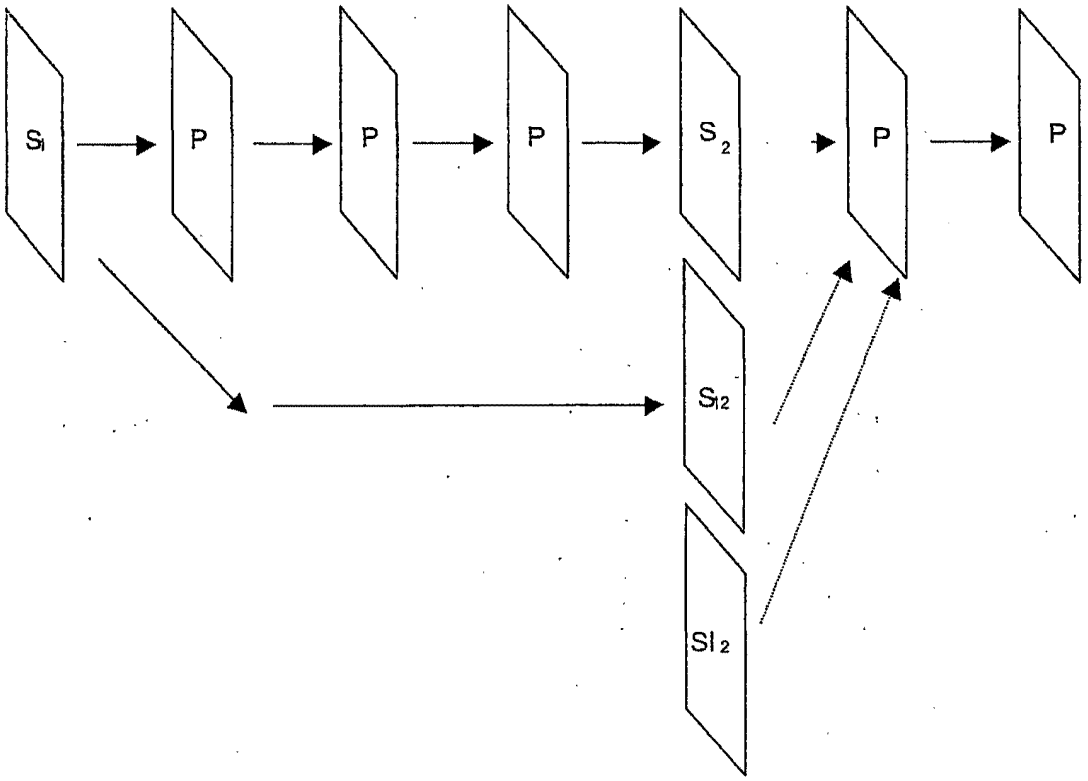


FIG. 14

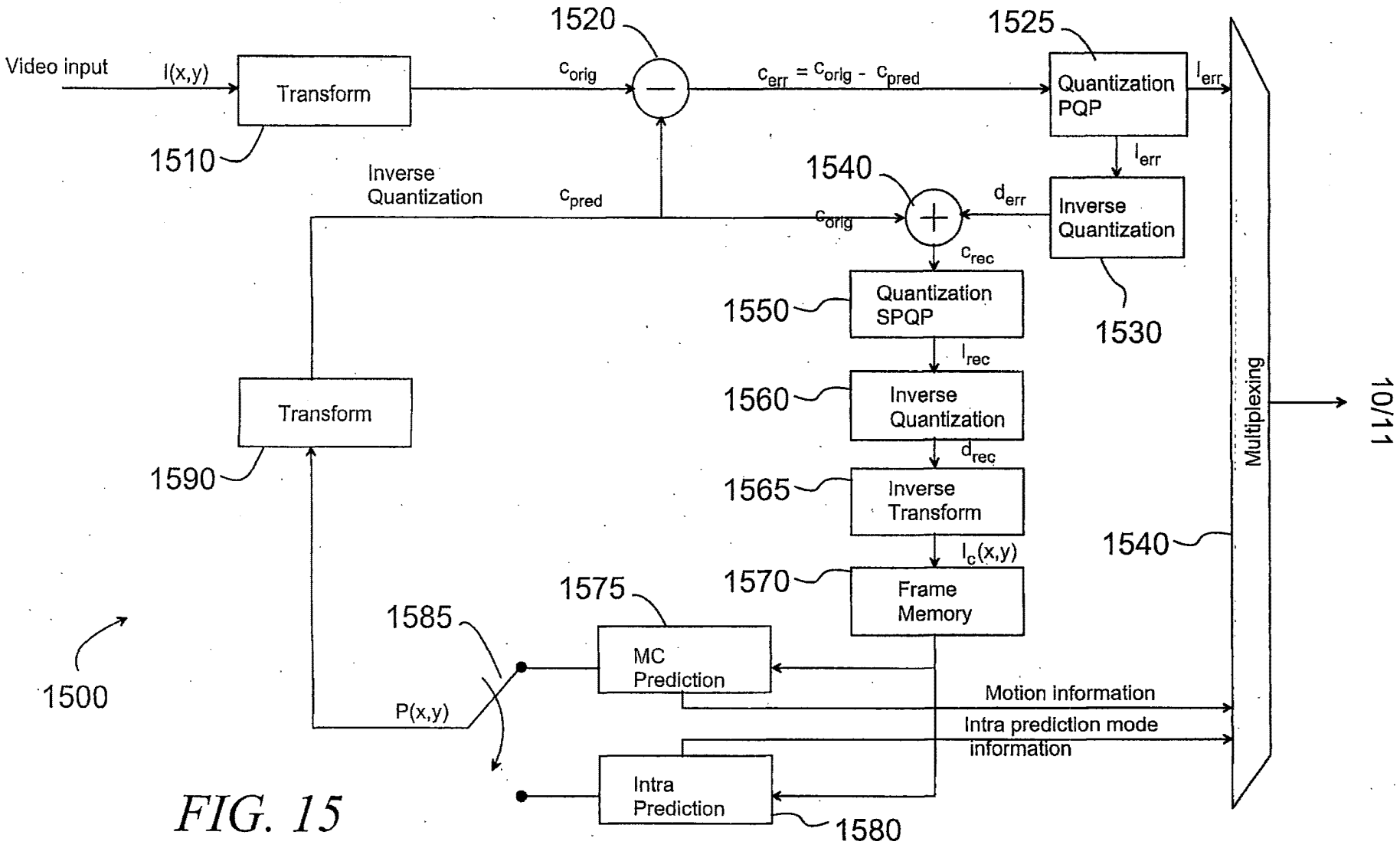

Video input
I(x,y)
Transform
1510
1520
c_{orig}
$c_{err} = c_{orig} - c_{pred}$
1525
Quantization PQP
l_{err}
l_{err}
Inverse Quantization
c_{pred}
1540
d_{err}
Inverse Quantization
c_{orig}
c_{rec}
1550
Quantization SPQP
1530
l_{rec}
1560
Inverse Quantization
Transform
1590
d_{rec}
1565
Inverse Transform
$I_c(x,y)$
1570
Frame Memory
1575
1585
MC Prediction
1540
Multiplexing
10/11
1500
P(x,y)
Motion information
Intra prediction mode information
Intra Prediction
1580

FIG. 15

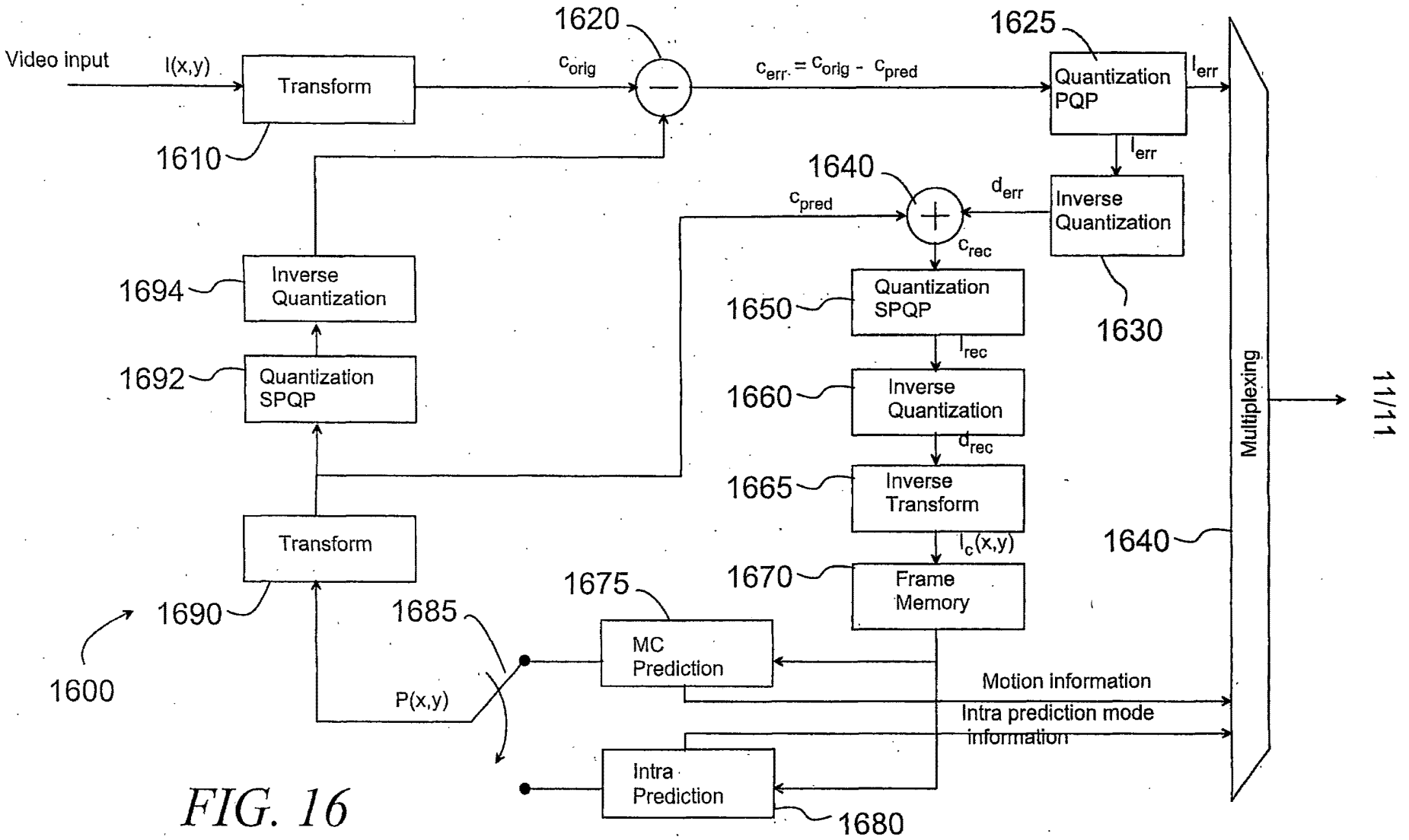

Video input
I(x,y)
Transform
1610
1620
c_{orig}
$c_{err} = c_{orig} - c_{pred}$
1625
Quantization PQP
I_{err}
I_{err}
1640
c_{pred}
d_{err}
Inverse Quantization
c_{rec}
1630
1694
Inverse Quantization
1692
Quantization SPQP
1650
Quantization SPQP
I_{rec}
1660
Inverse Quantization
d_{rec}
1665
Inverse Transform
$I_c(x,y)$
1670
Frame Memory
1640
Multiplexing
11/11
Transform
1690
1600
1675
1685
MC Prediction
P(x,y)
Motion information
Intra prediction mode information
Intra Prediction
1680

FIG. 16

INTERNATIONAL SEARCH REPORT

International application No.
PCT/FI 02/00004

A. CLASSIFICATION OF SUBJECT MATTER

IPC7: H04N 7/26, H04N 7/58
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC7: H04N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

SE,DK,FI,NO classes as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-INTERNAL, WPI DATA, INSPEC, COMPENDEX, TDB

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	EP 1045589 A2 (SONY UK LTD), 18 October 2000 (18.10.00), abstract ---	1-16
P,A	EP 1079631 A1 (SONY UK LTD), 28 February 2001 (28.02.01), abstract ---	1-16
A	WO 9114340 A1 (THOMSON CONSUMER ELECTRONICS), 19 Sept 1991 (19.09.91), whole document ---	1-16
A	Patent Abstracts of Japan, abstract of JP 12-165816 A (SONY CORP), 16 June 2000 (16.06.00), abstract --- --------	1-16

☐ Further documents are listed in the continuation of Box C. ☒ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
30 April 2002	02 -05- 2002
Name and mailing address of the ISA/ Swedish Patent Office Box 5055, S-102 42 STOCKHOLM Facsimile No. +46 8 666 02 86	Authorized officer Jesper Bergstrand/LR Telephone No. +46 8 782 25 00

Form PCT/ISA/210 (second sheet) (July 1998)

INTERNATIONAL SEARCH REPORT

Information on patent family members

28/01/02

International application No.
PCT/FI 02/00004

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 1045589 A2	18/10/00	GB 9908809 D	00/00/00
		JP 2001054110 A	23/02/01
EP 1079631 A1	28/02/01	GB 2353653 A	28/02/01
		GB 9920274 D	00/00/00
		JP 2001119305 A	27/04/01
WO 9114340 A1	19/09/91	AT 154485 T	15/06/97
		AU 7480691 A	10/10/91
		DE 69126525 D,T	25/09/97
		EP 0519995 A,B	30/12/92
		JP 2891773 B	17/05/99
		JP 5505080 T	29/07/93
		KR 235815 B	15/12/99

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 005 232 A3

(12) EUROPEAN PATENT APPLICATION

(88) Date of publication A3:
10.12.2003 Bulletin 2003/50

(43) Date of publication A2:
31.05.2000 Bulletin 2000/22

(21) Application number: **99120843.0**

(22) Date of filing: **25.10.1999**

(51) Int Cl.[7]: **H04N 7/26**, H04N 7/50, H04N 7/58, H04N 7/46

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE
Designated Extension States:
AL LT LV MK RO SI

(30) Priority: **24.11.1998 US 198867**

(71) Applicant: **General Instrument Corporation**
Horsham, Pennsylvania 19044 (US)

(72) Inventors:
- **Wang, Limin**
  San Diego, California 92128 (US)
- **Luthra, Ajay**
  San Diego, California 92130 (US)
- **Eifrig, Robert O.**
  San Diego, California 92126 (US)

(74) Representative: **Hoeger, Stellrecht & Partner**
Uhlandstrasse 14 c
70182 Stuttgart (DE)

(54) Rate control for an MPEG transcoder without a priori knowledge of picture type

(57) A rate control system suitable for use with a digital video transcoder, such as one conforming to the MPEG standard. The proposed rate control system starts coding without a priori knowledge of the Group of Pictures (GOP) parameters, thereby avoiding a processing delay of about one GOP which would otherwise be incurred to extract the complete GOP structure information. In addition, the system avoids the need to store the data corresponding to the GOP, thereby reducing the memory required for transcoding. An initially allocated bit budget is gradually corrected as successive pictures are coded according to their picture types. Changes in the GOP structure of pre-compressed bitstreams can be addressed, for example, when switching channels or inserting commercials. Target rates with incorrect starting GOP parameters will converge within a few GOPs.

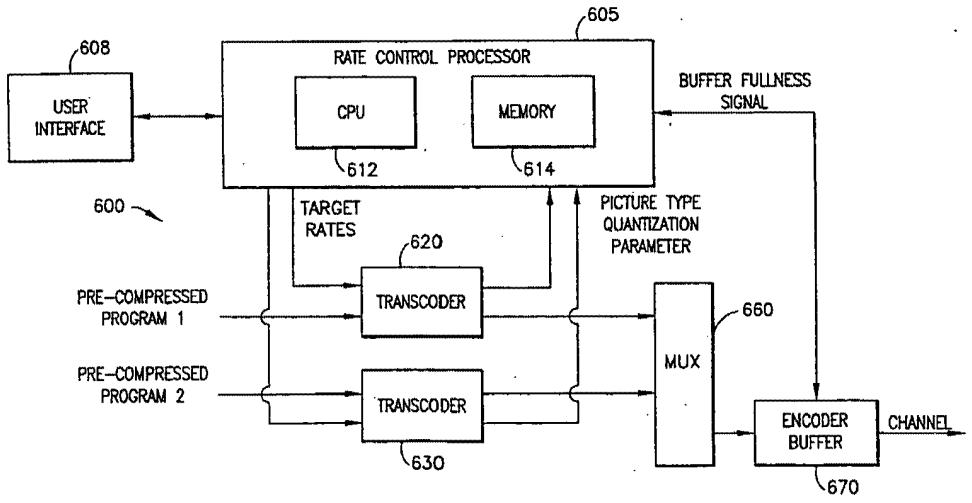


FIG.6

EP 1 005 232 A3

European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 99 12 0843

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.7)
Y	WILKINSON J: "UNDERSTANDING MPEG CONCATENATION AND TRANSCODING" ABU TECHNICAL REVIEW, ASIAN BROADCASTING UNION, KUALA LUMPUR, MW, no. 177, July 1998 (1998-07), pages 3-9, XP000992282 ISSN: 0126-6209 * section "TRANSCODING" * ---	1-3, 12-14, 17-19, 27,34, 36-38, 47-49, 52-54, 62,69	H04N7/26 H04N7/50 H04N7/58 H04N7/46
Y	LIANG-WEL LEE ET AL: "ON THE ERROR DISTRIBUTION AND SCENE CHANGE FOR THE BIT RATE CONTROLOF MPEG" IEEE TRANSACTIONS ON CONSUMER ELECTRONICS, IEEE INC. NEW YORK, US, vol. 39, no. 3, 1 August 1993 (1993-08-01), pages 545-554, XP000396331 ISSN: 0098-3063 * sections II., III. * ---	1-3, 12-14, 17-19, 27,34, 36-38, 47-49, 52-54, 62,69	
A	WO 98 44737 A (KONINKL PHILIPS ELECTRONICS NV ;PHILIPS NORDEN AB (SE)) 8 October 1998 (1998-10-08) * page 2, line 32 - page 4, line 25 * * page 14, line 21 - page 18, line 6 * ---	1-76	TECHNICAL FIELDS SEARCHED (Int.Cl.7) H04N
A	WO 95 29559 A (THOMSON CONSUMER ELECTRONICS ;OZKAN MEHMET KEMAL (US); BEYERS BILL) 2 November 1995 (1995-11-02) * page 10, line 35 - page 11, line 14 * * page 16, line 17 - page 17, line 27 * ---	1-76	
	-/--		

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
MUNICH	13 October 2003	Fassnacht, C

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 99 12 0843

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.7)
A	EP 0 700 214 A (HUGHES AIRCRAFT CO) 6 March 1996 (1996-03-06) * column 8, line 8 - column 8, line 31 *	1-3, 12-14, 17-19, 27, 34-38, 47-49, 52-54, 62,69	
A	PEREIRA M ET AL: "Re-codable video" PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON IMAGE PROCESSING (ICIP) AUSTIN, NOV. 13 - 16, 1994, LOS ALAMITOS, IEEE COMP. SOC. PRESS, US, vol. 3 CONF. 1, 13 November 1994 (1994-11-13), pages 952-956, XP010146085 ISBN: 0-8186-6952-7 * section "2.3.3. Improved rate-control" *	1-3, 12-14, 17-19, 27, 34-38, 47-49, 52-54, 62,69	
P,X	LIMIN WANG ET AL: "Adaptive rate control for MPEG transcoder" IMAGE PROCESSING, 1999. ICIP 99. PROCEEDINGS. 1999 INTERNATIONAL CONFERENCE ON KOBE, JAPAN 24-28 OCT. 1999, PISCATAWAY, NJ, USA,IEEE, US, 24 October 1999 (1999-10-24), pages 266-270, XP010368706 ISBN: 0-7803-5467-2 * the whole document *	1-76	TECHNICAL FIELDS SEARCHED (Int.Cl.7)
E	WO 01 84905 A (GEN INSTRUMENT CORP ;WANG LIMIN (US)) 15 November 2001 (2001-11-15) * the whole document *	1-76	

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
MUNICH	13 October 2003	Fassnacht, C

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

ANNEX TO THE EUROPEAN SEARCH REPORT ON EUROPEAN PATENT APPLICATION NO. EP 99 12 0843

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report. The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

13-10-2003

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9844737 A	08-10-1998	US 5982436 A	09-11-1999
		EP 0914745 A1	12-05-1999
		WO 9844737 A1	08-10-1998
		JP 2000511740 T	05-09-2000
		US 6208691 B1	27-03-2001
WO 9529559 A	02-11-1995	WO 9529559 A1	02-11-1995
		AU 6816094 A	16-11-1995
		EP 0803163 A1	29-10-1997
		JP 9512396 T	09-12-1997
		US 6055270 A	25-04-2000
EP 0700214 A	06-03-1996	US 5708473 A	13-01-1998
		EP 0700214 A2	06-03-1996
		JP 3072035 B2	31-07-2000
		JP 8205180 A	09-08-1996
		NO 953369 A	01-03-1996
WO 0184905 A	15-11-2001	AU 6111901 A	20-11-2001
		CA 2407682 A1	15-11-2001
		CN 1433641 T	30-07-2003
		EP 1295462 A1	26-03-2003
		TW 522740 B	01-03-2003
		WO 0184905 A2	15-11-2001

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 005 232 A2

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
31.05.2000 Bulletin 2000/22

(51) Int. Cl.[7]: **H04N 7/26**, H04N 7/50, H04N 7/58

(21) Application number: **99120843.0**

(22) Date of filing: **25.10.1999**

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE
Designated Extension States:
AL LT LV MK RO SI

(30) Priority: **24.11.1998 US 198867**

(71) Applicant: **General Instrument Corporation**
Horsham, Pennsylvania 19044 (US)

(72) Inventors:
, Wang, Limin
San Diego, California 92128 (US)
, Luthra, Ajay
San Diego, California 92130 (US)
, Eifrig, Robert O.
San Diego, California 92126 (US)

(74) Representative:
Hoeger, Stellrecht & Partner
Uhlandstrasse 14 c
70182 Stuttgart (DE)

(54) Rate control for an MPEG transcoder without a priori knowledge of picture type

(57) A rate control system suitable for use with a digital video transcoder, such as one conforming to the MPEG standard. The proposed rate control system starts coding with any reasonable set of assumed Group of Pictures (GOP) parameters, thereby avoiding a processing delay of about one GOP which would otherwise be incurred to extract the complete GOP structure information from a pre-compressed bit stream. In addition, the system avoids the need to store the data corresponding to the GOP, thereby reducing the memory required for transcoding. Encoding of a first picture in a sequence or GOP begins without *a priori* knowledge of the picture type of subsequent pictures. A reasonable set of GOP parameters is assumed to determine an encoding bit budget. The bit budget is gradually corrected as successive pictures are coded according to their picture types. Changes in the GOP structure of pre-compressed bitstreams can be addressed, for example, when switching channels, inserting commercials, and the like. Target rates with incorrect starting GOP parameters will converge within a few GOPs.

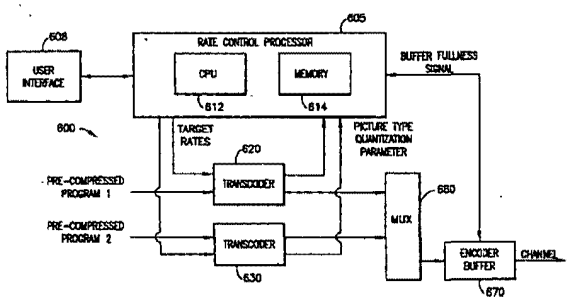


FIG.6

EP 1 005 232 A2

Description

BACKGROUND OF THE INVENTION

[0001] The present invention relates to a digital video transcoder and method for allocating bits for encoding successive pictures in a group of pictures (GOP) without *a priori* knowledge of the picture types in the GOP.

[0002] With digital video coding standards such as MPEG, input pictures can be coded in three different picture types, namely I, P and B. The three pictures require quite different numbers of bits for encoding because of the different nature of their temporal processing. Hence, an intelligent bit allocation strategy should assign a number of bits for encoding according to the picture's type. This implies a requirement of *a priori* knowledge of the picture types for a given bit budget. This requirement is not a problem for a standalone encoder as the encoder can determine the picture type for each input picture.

[0003] In fact, the encoder can plan ahead for the types of the input pictures. In contrast, a transcoder has no such *a priori* knowledge regarding a picture's type before actually processing the picture. This creates difficulty in allocating an appropriate number of bits for encoding pictures in a transcoder.

[0004] Accordingly, it would be desirable to have a method and apparatus for allocating bits for encoding pictures in a transcoder without *a priori* knowledge of the picture type. The system should avoid a processing delay of about one GOP which would otherwise be incurred to extract the complete GOP structure information from a pre-compressed bit stream. In addition, the system should avoid the need to store the data corresponding to the GOP, thereby reducing the memory required for transcoding.

[0005] The system should be compatible with both variable bit rate and constant bit rate bitstreams.

[0006] The system should be compatible with statistical multiplexing and remultiplexing systems.

[0007] The present invention provides an apparatus and method having the above and other advantages.

SUMMARY OF THE INVENTION

[0008] The present invention relates to a digital video transcoder and method for allocating bits for encoding successive pictures in a group of pictures (GOP) without *a priori* knowledge of the arrangement of picture types in the GOP.

1. Progressive Refresh Sequence:

[0009] A method for allocating bits for coding a progressive refresh sequence of pictures in a bitstream received at a digital video transcoder includes the steps of: (a) providing an assumed distance M' between a first picture of the bitstream and the next closest subsequent P-picture of the bitstream (in display order); (b) providing an assumed bit budget as a function of the assumed distance M'; (c) coding the first picture in accordance with the assumed bit budget; (d) determining a picture type of a second picture that immediately follows the first picture in the bitstream; (e) adjusting the assumed bit budget according to the picture type of the second picture; and (f) allocating bits for coding the second picture in accordance with the adjusted bit budget.

[0010] Thus, a bit budget for coding the pictures is initially assumed for the first picture, and the bit budget is updated as the second picture type is known. A target number of bits is allocated for coding each picture according to the bit budget. The first and second picture types generally indicate the picture distribution in the sequence.

[0011] The assumed bit budget is proportional to the assumed distance and a bit rate of the bitstream, and inversely proportional to a frame rate of the bitstream.

[0012] The pictures in the bitstream may form a progressive refresh sequence, where there is no GOP structure.

1.1 P-picture followed by B-picture:

[0013] When the first picture is a P-picture and the second picture is a B-picture, the method includes the further step of: determining an actual distance M" between the first picture and the next closest subsequent P-picture (in display order) according to the picture type of the first picture and the picture type of the second picture.

[0014] The actual distance M" is determined according to a difference between a temporal reference of the first picture, in a display order of the bitstream, and a temporal reference of the second picture, in the display order, plus one picture.

[0015] The method may include the further step of allocating bits for coding the remaining M"-1 pictures following the first picture in accordance with the adjusted bit budget.

[0016] The assumed bit budget is adjusted in the adjusting step (f) by +(M"-M')*"bit rate"/"frame rate", where "bit rate" is a bit rate of the bitstream, and "frame rate" is a frame rate of the bitstream.

1.2 P-picture followed by P-picture:

[0017] When the first and second pictures are P-pictures, the method includes the further step of: determining an actual distance M" between the first picture and the next closest subsequent P-picture (in display order)

according to the picture type of the first picture and the picture type of the second picture. The actual distance M" is determined according to a difference between a temporal reference of the second picture and a temporal reference of the first picture. The temporal references are determined in relation to a display order of the bitstream.

[0018] The assumed bit budget is adjusted in the adjusting step (f) by -(M'-1)*"bit rate"/"frame rate", where "bit rate" is a bit rate of the bitstream, and "frame rate" is a frame rate of the bitstream.

[0019] The method includes the further step of allocating bits for coding a series of M" pictures following an initial M" pictures that includes the first and second pictures in the bitstream in accordance with the adjusted bit budget. Thus, the adjustments to the bit budget in one series of pictures is carried over to the next series to allow correct coding of the next series.

2. Non-progressive Refresh Sequence:

[0020] Generally, the invention enables transcoding to begin by assuming a reasonable set of GOP parameters, including M, the distance between each I-picture and the next P-picture, and N, the GOP length. M and N are adjusted during transcoding as additional information becomes available regarding the GOP structure, e.g., the distribution of picture types in the GOP and the length of the GOP.

[0021] The assumed M can be corrected to the actual value within two frames, and the assumed N can be corrected to the actual value within one GOP. M should be verified, and adjusted if necessary, after each I- or P-picture, and N should be verified, and adjusted if necessary, after each I-picture. When M and/or N are adjusted, a bit rate R for coding the pictures, and n_I, n_P, and n_B, the remaining numbers of I, P and B-pictures in the current GOP, respectively, are also adjusted.

[0022] A method for allocating bits for coding a non-progressive refresh sequence of pictures in a bitstream received at a digital video transcoder, includes the steps of: (a) providing an assumed length N' of a particular group of pictures (GOP) of the bitstream; (b) providing an assumed bit budget as a function of the assumed length N'; providing an assumed distance M' between the first picture and the next closest subsequent P-picture of the bitstream (in display order); (c) allocating bits for coding a first picture of the particular GOP in accordance with the assumed bit budget, and N' and M'; (d) determining a picture type of a second picture that immediately follows the first picture; (e) adjusting the assumed bit budget according to the picture type of the second picture; and (f) allocating bits for coding the second picture in accordance with the adjusted bit budget.

[0023] The assumed bit budget is proportional to the assumed length N' and a bit rate of the bitstream, and inversely proportional to a frame rate of the bitstream.

[0024] M" may change within a GOP. Thus, the method may include the step of periodically verifying M" throughout the GOP. For example, M" may be calculated at every I-picture and/or P-picture that follows the first I-picture. The assumed bit budget is adjusted each time in accordance with M".

2.1 I-picture followed by I-picture:

[0025] When the first and second pictures are I-pictures, the method comprises the further step of: determining an actual distance M" between the first picture and the next closest subsequent I-picture of the bitstream according to the picture type of the second picture, and adjusting the assumed bit budget in accordance with M"

[0026] The actual distance M" is determined according to a difference between a temporal reference of the second picture, in a display order of the bitstream, and a temporal reference of the first picture, in the display order.

[0027] The adjusting step (e) comprises the step of: adjusting the assumed bit budget by -(N'-1)*"bit rate"/"frame rate", where "bit rate" is a bit rate of the bitstream, "frame rate" is a frame rate of the bitstream, and N=M=M".

[0028] The method comprises the further step of: allocating bits for coding the remaining pictures in the particular GOP following the second picture in accordance with the adjusted bit budget.

2.2 I-picture followed by P-picture:

[0029] When the first picture is an I-picture and the second picture is a P-picture, the method comprises the further steps of: determining an actual distance M" between the first picture and the next closest subsequent P-picture (in display order) according to the picture type of the second picture; and adjusting the assumed length N' of the GOP in accordance with the actual distance M" to provide an adjusted assumed length N"; wherein: the adjusting step (e) is responsive to the adjusted assumed length N'.

[0030] The assumed length N' of the particular GOP is adjusted in the adjusting step thereof by a factor, M(N'/M"), where M=M", to provide the adjusted assumed length N".

[0031] The adjusting step (e) comprises the step of: adjusting the assumed bit budget by a factor (N"- N')*"bit rate"/"frame rate", where "bit rate" is a bit rate of the bitstream, and "frame rate" is a frame rate of the bitstream.

[0032] The adjusting step (e) comprises the still further step of: adjusting the assumed bit budget by a factor, (M"-1')*"bit rate"/"frame rate", where "bit rate" is a bit rate of the bitstream, "frame rate" is a frame rate of the bitstream, and M=M".

[0033] Thus, both the GOP length and the distance M are used to adjust the bit budget, which in turn affects the target bit allocation for each picture.

[0034] The method includes the further step of allocating bits for coding the remaining pictures in the particular GOP following the second picture in accordance with the adjusted bit budget.

[0035] The method includes the further steps of determining if N">N or N"<N, where N is an actual length of the current GOP, after coding N" pictures in the current GOP. This is accomplished by observing the picture that is reached after coding N" pictures. For example, if the picture reached after coding N" pictures is a P-picture in the current GOP, then N"<N. If the picture reached after coding N" pictures is an I-picture, this signals the start of the next GOP, so N"=N. If an I-picture is coded before the N" are all coded, this indicates N">N. The already adjusted bit budget is then further adjusted based on whether N"<N or N">N, and bits are allocated for coding pictures in the bitstream following the N" coded pictures according to the further adjusted bit budget.

2.3 I-picture followed by B-picture:

[0036] When the first picture is an I-picture and the second picture is a B-picture, the method comprises the further steps of: determining an actual distance M" between the first picture and the next closest subsequent P-picture (in display order) according to the picture type of the second picture; and adjusting the assumed length N' of the GOP in accordance with the actual distance M" to provide an adjusted assumed length N"; wherein: the adjusting step (e) is responsive to the adjusted assumed length N'.

[0037] The assumed length N' of the particular GOP is adjusted in the adjusting step thereof by a factor, M(N'/M''), where M=M", to provide the adjusted assumed length N".

[0038] The adjusting step (e) comprises the step of: adjusting the assumed bit budget by a factor (N"- N')*"bit rate"/"frame rate", where "bit rate" is a bit rate of the bitstream, and "frame rate" is a frame rate of the bitstream.

[0039] The method comprises the further steps of: allocating bits for coding the remaining pictures in the particular GOP following the second picture in accordance with the adjusted bit budget.

[0040] Additionally, as discussed above, the already adjusted bit budget may be further adjusted based on whether N"<N or N">N, and bits may be allocated for coding pictures in the bitstream following the N" coded pictures according to the further adjusted bit budget.

[0041] Thus, the information learned regarding GOP length and M, the distance between I and/or P-pictures, in one GOP are used in coding the following GOPs. Target rates with incorrect starting GOP parameters will converge within a few GOPs.

[0042] Corresponding apparatus structures are also presented.

BRIEF DESCRIPTION OF THE DRAWINGS

[0043]

FIG. 1(a) illustrates a progressive refresh sequence of pictures in display order for continuing GOPs.

FIG. 1(b) illustrates the sequence of pictures of FIG. 1(a) in encoding order.

FIG. 2(a) illustrates a progressive refresh sequence of pictures in display order for an initial GOP in a bitstream followed by continuing GOPs.

FIG. 2(b) illustrates the sequence of pictures of FIG. 2(a) in encoding order.

FIG. 3 illustrates an MPEG-2 bitstream structure.

FIG. 4(a) illustrates a coding method for a progressive refresh picture sequence in accordance with the present invention.

FIG. 4(b) illustrates a coding method for a non-progressive refresh picture sequence in accordance with the present invention.

FIG. 4(c) illustrates a continuation of the coding method of FIG. 4(b) in accordance with the present invention.

FIG. 5 illustrates an estimated GOP length N" in accordance with the present invention.

FIG. 6 illustrates a rate control system in accordance with the present invention.

FIGs 7(a)-7(d) illustrate a C-language flowchart for implementing the present invention.

DETAILED DESCRIPTION OF THE INVENTION

[0044] The present invention relates to a digital video transcoder and method for allocating bits for encoding

successive pictures in a group of pictures (GOP) without *a priori* knowledge of the picture types in the GOP.

[0045] A novel rate control scheme and apparatus is presented for a digital video transcoder, such as one conforming to the MPEG standard, that requires no *a priori* (e.g., beforehand) knowledge of the picture types. Studies indicate that the picture target rates determined by the proposed rate control system with and without *a priori* knowledge of the picture types are very close.

[0046] MPEG defines three picture types in terms of temporal processing. They are the intra-frame coded picture (I-picture), forward temporal predictive coded picture (P-picture), and the bi-directional temporal predictive coded picture (B-picture). In MPEG video coding, pictures of an input video sequence are often grouped into groups of pictures (GOPs), where each GOP may contain one I-picture, a number of P-pictures, and optionally one or more B-pictures between I or P-pictures. The structure of a GOP can often be described by two variables: (1)N - the number of pictures in a GOP, or the GOP length, and (2) M - the distance between I or P-pictures. An example of a GOP with N = 15 and M = 3, in display order, is:

$$\cdots B_{m-2}B_{m-1}I_mB_{m+1}B_{m+2}P_{m+3}B_{m+4}B_{m+5}P_{m+6}\cdots P_{m+N-3}B_{m+N-2}B_{m+N-1}I_{m+N}B_{m+N+1}\cdots$$

where the subscripts are the picture's temporal references, and the consecutive numbers denote consecutively displayed pictures.

[0047] In general, I-pictures require many more bits for encoding than P- and B-pictures because I-pictures do not take advantage of temporal correlations between successive pictures (e.g., frames). B-pictures use the least numbers of bits for encoding because they have two temporal references and can therefore be temporally predicted more efficiently than P-pictures. P-pictures use a number of bits for encoding that is generally between that used for B- and I-pictures.

[0048] Clearly, the bits assigned to each picture for encoding should be based on its type, or its need. Furthermore, a GOP can have all the three types of pictures, and the organization of the three picture types within a GOP can be very flexible. Therefore, to wisely allocate a bit budget for encoding the pictures within a GOP, it is necessary to have *a priori* knowledge of the types of pictures in the GOP, or the GOP structure. This requirement is not a problem for a standalone encoder (i.e., an encoder that is not part of a transcoder) since the encoder can decide the type for each picture, and plan ahead in allocating bits based on the type and arrangement of pictures in a GOP. In other words, the information on GOP structure is available to the encoder.

[0049] In contrast to the standalone encoder, a transcoder converts a pre-compressed bitstream into another bitstream at a new rate. The inputs to a transcoder are pre-compressed bitstreams. Before processing a picture, the transcoder has no *a priori* knowledge of the picture's type, and hence no *a priori* knowledge of the GOP structures of the pre-compressed bitstreams.

[0050] Moreover, some transcoding schemes re-use the motion vector (MV) fields of the received pre-compressed bitstream for encoding the output bitstream at the new data rate. In this case, the picture types of the pre-compressed pictures are maintained. This means that the transcoder has no control of the GOP structures of the new output bitstreams.

[0051] Furthermore, with conventional schemes, it is possible for the GOP structure information to be extracted from the pre-compressed bitstreams. For example, after scanning a pre-compressed bitstream by the first two I-pictures, a transcoder should be able to determine the GOP structure. However, this requires a processing delay of about one GOP and a requirement of extra memory to hold all the bits for the GOP.

[0052] The present invention discloses a novel rate control system for a digital video transcoder, such as that conforming to the MPEG standard, with no requirement of *a priori* knowledge of the GOP structures of the pre-compressed bitstreams. A rate control scheme is disclosed that causes no processing delay and requires no extra memory. The scheme starts with any reasonable set of GOP parameters, and then gradually corrects these GOP parameters, if necessary, as successive pre-compressed pictures are received at the transcoder. The scheme compensates for changing GOP parameters used by the encoder(s).

[0053] Furthermore, we analyze the differences in target bit rates determined by the proposed rate control scheme with and without *a priori* knowledge of the GOP structure of the pre-compressed bitstream. We show that the target rates with and without *a priori* knowledge of GOP structure tend to be the same within only one or two GOPs. We demonstrate the differences in target rates in both numerical examples and real video sequences.

[0054] Below, in section 1, we will first overview the rate control for MPEG video coding. In section 2, we present the novel rate control for MPEG transcoder. In section 3, we analyze the differences in target rates determined by the proposed rate control with and without *a priori* knowledge of GOP structure. In section 4, we present the simulation and comparison results. Conclusions are discussed in section 5.

1. RATE CONTROL FOR MPEG VIDEO CODING

[0055] In MPEG video coding, pictures of all the three types (I, P or B) undergo the block Discrete Cosine Transform (DCT). The DCT coefficients are then quantized and variable length encoded. The resulting number of bits for each picture depends on the complexity of its corresponding block DCT image. I-pictures are coded without reference to any other pictures and, therefore, their DCT images usually require the most number of bits. P-pictures are first forward temporally predicted, and their temporal prediction difference images are then coded. The temporal prediction difference images usually contain much less information than the original pictures, and therefore need less bits for coding. B-pictures with two temporal references can often be better temporally predicted than P-pictures. Hence, B-pictures require the least number of bits. An intelligent bit allocation strategy should assign

the bits to a picture based on the picture's type, or the picture's complexity.

[0056] Let X_n be the complexity measure for picture n in a GOP of N pictures. In MPEG Test Model 5 (TM5), the complexity measure for a picture is defined as the product of the average quantization parameter used for the picture and the resulting number of bits. Logically, the bits allocated for a picture should be proportional to the picture's complexity measure, X_n, i.e.,

$$T_n = CX_n \qquad n = 0,1,\ldots,N-1 \qquad (1)$$

where C is a constant. Here, we have assumed a linear model between the picture complexity measure and the number of bits required for coding. Furthermore, without any bias, a GOP of N pictures should be allocated

$$R_{GOP,N} = N \cdot \frac{bits\_rate}{frame\_rate} \text{ bits.} \qquad (2)$$

[0057] Clearly, the total number of bits assigned to all the pictures within a GOP of N pictures should be equal to $R_{GOP,N}$, i.e.,

$$R_{GOP,N} = \sum_{n=0}^{N-1} T_n = \sum_{n=0}^{N-1} CX_n = C\sum_{n=0}^{N-1} X_n \qquad (3)$$

[0058] From equations (1) and (3), we have

$$T_n = \frac{X_n}{\sum_{n=0}^{N-1} X_n} R_{GOP,N} \qquad (4)$$

[0059] Assume that all the pictures of the same type, i.e., I, P or B, have the same complexity measure, which is a reasonable assumption, at least for continuous scenes. We only need three complexity measures, one for each picture type. That is, X_I for I-pictures, X_P for P-pictures, and X_B for B-pictures. The target rate for frame n of type $t \in \{I,P,B\}$ and complexity measure X_n, $t_{\in\{I,P,B\}}$ can therefore be written as:

$$T_n = \frac{X_{n,t}}{N_I X_I + N_P X_P + N_B X_B} R_{GOP,N} \qquad (5)$$

where N_I, N_P and N_B are the numbers of I, P and B-pictures, respectively, in the GOP. By eqn. (5), the pictures of the same type are assigned the same number of bits. This bit allocation strategy can be further modified by considering only the distribution of the remaining number of bits over the remaining pictures in the current GOP, i.e.,

$$T_n = \frac{X_{n,t}}{n_I X_I + n_P X_P + n_B X_B} R \qquad (6)$$

where n_I, n_P and n_B are, respectively, the remaining numbers of I, P and B-pictures in the current GOP, and R is the remaining number of bits, defined as

$$R \Leftarrow R - \sum_{n'=0}^{n-1} T_{n'} \qquad (7)$$

where $\Leftarrow$ denotes the assignment statement.

[0060] The remaining number of bits, R, needs to be updated at the beginning of each GOP as,

$$R = R + R_{GOP,N} = R + N \cdot \frac{bit\_rate}{frame\_rate} \qquad (8)$$

[0061] We now show that eqn. (6) also assigns the same number of bits to the pictures of the same type as long as the encoder can meet the target rate at each picture. For example, assume that frame n is a B-picture. Frame n can actually be either an I, P or B-picture. The encoder meets the nominal rate, T_n, for frame n. The target rate for frame $n+1$, Tn_{+1}, is therefore:

$$T_{n+1} = \frac{X_{n+1,t}}{n_I X_I + n_P X_P + (n_B - 1) X_B}(R - T_n) \qquad (9a)$$

$$= \frac{X_{n+1,t}}{n_I X_I + n_P X_P + (n_B - 1) X_B}\left(R - \frac{X_B}{n_I X_I + n_P X_P + n_B X_B} R\right) \qquad (9b)$$

$$= \frac{X_{n+1,t}}{n_I X_I + n_P X_P + (n_B - 1) X_B}\left(\frac{n_I X_I + n_P X_P + n_B X_B - X_B}{n_I X_I + n_P X_P + n_B X_B}\right) R \qquad (9c)$$

$$= \frac{X_{n+1,t}}{n_I X_I + n_P X_P + n_B X_B} R \qquad (9d)$$

[0062] As seen, the target rate for frame n+1, Tn_{+1}, depends upon only the type of frame n+1, i.e., $Xn_{+1}t$. In other words, regardless of the type (either I, P or B) of frame n, frame n+1 will be assigned the same number of bits according to its own type. This implies that equations (5) and (6) are actually the same if the pictures of the same type have the same complexity measure, and the encoder can meet the target rate at each picture. Equation (6) is, however, more practical as it addresses the changes in picture complexity measures and the difference between the target and actual rates at each picture.

[0063] In practice, the quality requirement may not be the same for the different picture types. For example, B-pictures are never used as references in the future temporal prediction. Hence, they can tolerate more distortion than I- and P-pictures. To address this concern of different quality requirements, we can further introduce a weighting factor for each picture type, $K_{n,t}$, as

$$T_n = CX_{n,t} / K_{n,t} \qquad (10)$$

where

$$K_t = \begin{cases} K_I & for\ I\ picture \\ K_P & for\ P\ picture \\ K_B & for\ B\ picture \end{cases} \qquad (11)$$

[0064] The bit allocation strategies (e.g., equations 5 and 6) now become:

$$T_n = \frac{X_{n,t}/K_{n,t}}{N_I X_I / K_I + N_P X_P / K_P + N_B X_B / K_B} R_{GOP,N} \qquad (5b)$$

and

$$T_n = \frac{X_{n,t}/K_{n,t}}{n_I X_I / K_I + n_P X_P / K_P + n_B X_B / K_B} R \qquad (6b)$$

[0065] Equation (6b) is the same one used in MPEG TM5 when $K_I = K_P = 1$ and $K_B = 1.4$.

[0066] The bit allocation strategy (e.g., equations 5 and 6) requires *a priori* knowledge of GOP structure, i.e., the numbers of I-, P- and B-pictures in a GOP, N_I,N_P,and N_B, or the remaining number of I, P and B-picture in the current GOP, n_I,n_P and n_B Given a pair of GOP structure variables, N and M, we can have $N_I = 1$, $N_P = N/M-N_I$ and $N_B = N-N_P-N_I$. FromN_I,N_P,and N_B, we can also calculate n_I,n_P and n_B during encoding. For example, if a P-picture is just encoded, the remaining number of P-pictures in the current GOP is decreased by 1, i.e., $n_P = n_P$-1.

[0067] As discussed, this information is unfortunately not available for a transcoder. Without a complete knowledge of the GOP structure of pre-compressed bitstreams, it is difficult to perform an intelligent bit allocation strategy. A transcoder can learn the GOP structure of a pre-compressed bitstream by scanning the bitstream through the first two I-pictures. This, however, may cause a processing delay by one GOP and require extra memory for holding the bits of the GOP.

2. RATE CONTROL FOR A TRANSCODER

[0068] We now describe a novel rate control system for a transcoder with no requirement of *a priori* knowledge of GOP structure of the pre-compressed bitstreams. The proposed rate control system causes no processing delay and requires no extra memory. It can start with any reasonable set of GOP parameters, and then gradually corrects these GOP parameters as encoding progresses, if necessary.

[0069] Table 1(a) shows all the possible picture organizations where the subscripts are the pictures' temporal references, in display order. Each picture refers to a frame or field, or portion thereof. For example, the present invention is compatible with coding of video object planes (VOPs) as known from the MPEG-4 standard. The present invention is compatible with both progressive scan and interlaced scan video.

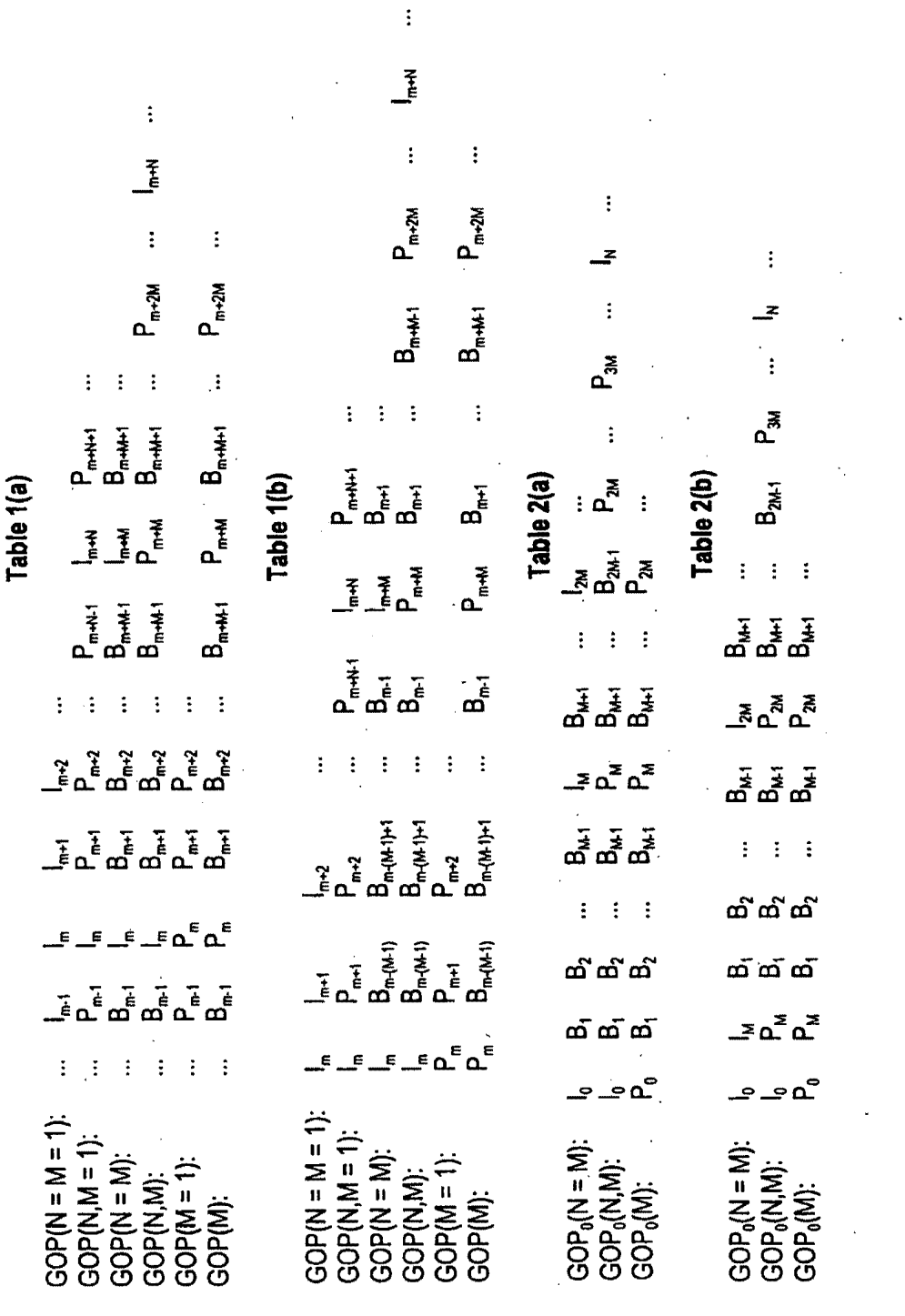
[0070] For the sequence GOP(N=M=1), there are only I-pictures. For the sequence GOP(N,M=1), there are only I- and P-pictures. For the sequence GOP(N=M), there are only I- and B-pictures. For the sequence GOP(N,M), there are I-, P- and B-pictures. For the sequence GOP(M=1), there are only P-pictures. For the sequence GOP(M), there are only P- and B-pictures.

[0071] The first four cases in Table 1(a) are non-progressive refresh sequences since they have I-pictures. The last two case (e.g., GOP(M=1) and GOP(M)) have no I-pictures, but use a progressive refresh where P-pictures contain I-slices. An I-slice is an intra-coded portion of a picture typically extending horizontally across the picture. The location of the I-slice changes in successive P-pictures to refresh a different portion of the successive P-pictures. Hence, with a progressive refresh, the GOP length N is no longer necessary in describing the picture organization. Nevertheless, M can still be used to indicate the distance between two consecutive P-pictures. The "distance" is measured in terms of number of pictures herein, although any corresponding measure, such as time, can be used. In terms of bit allocation, we can set $N = M$, $N_I = 0$ (as there is no I pictures for progressive refresh sequence), $N_P = N/M - N_I = 1$, $N_B = N - N_P - N_I = N - N_P$ in $N = M$ bit allocation equation (5) for progressive refresh.

[0072] Table 1(b) shows the corresponding forms of picture organizations in encoding order (e.g., in bitstream order) for the picture sequences of Table 1(a). If $M = 1$ (there is no B-picture in the sequence), the encoding order is actually the same as the display order. On the other hand, if $M \neq 1$ (e.g., there are M -1 B-pictures between successive I- or P-pictures), the future I- or P-pictures have to be encoded before the M -1 B-pictures.

[0073] Furthermore, for the first GOP (e.g., initial or non-continuous GOP, designated GOP_0) in a video sequence with $M \neq 1$, there may not be any B-pictures before the first I or P-picture, as shown in Table 2(a). Table 2(a) shows possible picture organizations of the first (initial) GOP after a sequence header in a video stream, in display order. The initial GOPs are designated GOP_0. For example, for GOP_0(M), the first picture is designated P_0. An initial GOP may be processed by a transcoder when a bitstream is initially acquired or re-acquired, for example.

[0074] Table 2(b) shows the first GOP of Table 2(a) in encoding order. Here, an initial GOP with $M \neq 1$ is shown, where the second I- or P-picture is encoded before any B-picture. The second picture in encoding order will be either an I or P-picture, as shown in Table 2(b).



[0075] FIG. 1(a) illustrates a progressive refresh sequence of pictures in display order for continuing GOPs. Specifically, the progressive refresh sequence GOP(M) from Table 1(a) is illustrated for a distance between P-pictures of M=3, and for a continuing (e.g., non-initial, or open) GOP. Additionally, the GOP length is N=M=3 pictures.

[0076] In the video sequence 100, a first GOP 110 includes pictures P_7, B_8 and B_9, while a second GOP 120 includes pictures P_{10}, B_{11} and B_{12}, and a third GOP 130 includes pictures P_{13}, B_{14}, and B_{15}. The subscript "7" for P_7 is chosen simply as an example. A subsequent GOP includes P_{16} as its first picture.

[0077] FIG. 1(b) illustrates the sequence of pictures of FIG. 1(a) in encoding order. The encoded video sequence

150 corresponds to the sequence GOP(M) from Table 1(b). Since B_8 and B_9 are temporally backward predicted from P_{10}, B_8 and B_9 follow P_{10} in the encoded sequence 150. Similarly, B_{11} and B_{12} follow P_{13}, and B_{14} and B_{15} follow P_{16}.

[0078] FIG. 2(a) illustrates a progressive refresh sequence of pictures in display order for an initial GOP in a bitstream followed by continuing GOPs. Specifically, the progressive refresh sequence $GOP_0(M)$ from Table 2(a) is illustrated for a distance between P-pictures of M=3, and a GOP length is N=M=3 pictures.

[0079] In the video sequence 200, a first, initial GOP 210 includes pictures P_0, B_1 and B_2, while a second, continuing GOP 220 includes pictures P_3, B_4 and B_5, and a third, continuing GOP 230 includes pictures P_6, B_7, and B_8. A subsequent GOP includes P_9 as its first picture.

[0080] FIG. 2(b) illustrates the sequence of pictures of FIG. 2(a) in encoding order. The encoded video sequence 250 corresponds to the sequence $GOP_0(M)$ from Table 2(b). Since B_1 and B_2 are temporally backward predicted from P_3, B_1 and B_2 follow P_3 in the encoded sequence 250. Essentially, the B-pictures that would use P_0 for backward temporal prediction are not present, so P_3, B_1 and B_2 follow P_0. Similarly, B_4 and B_5 follow P_6, and B_7 and B_8 follow P_9.

[0081] FIG. 3 illustrates an MPEG-2 bitstream structure. A sequence layer 300 includes a sequence header 302, followed by a number of GOPs, e.g., GOPs 304, 306 and 308, followed by additional sequence headers and GOPs, including sequence header 312. A GOP layer 320 includes a GOP header 322, followed by data from a number of pictures, e.g., pictures 324, 326 and 328. A picture layer 330 includes a picture header 332, followed by a number of slices, e.g., slices 334 and 336. A slice layer 335 includes a slice header 337, followed by a number of macroblocks, e.g., macroblocks 338 and 339. A macroblock layer 340 includes a macroblock header 342, followed by a number of blocks, e.g., blocks 344, 346 and 348. A block layer includes DCT coefficients 352.

[0082] For a given pre-compressed bitstream, the transcoder can start processing anywhere in the bitstream. Initially, the transcoder will skip all the bits until reaching the sequence header 300, which contains information about picture size, frame rate, and the like. Optionally, the GOP header 322 may follow the sequence header 302 (for non-progressive refresh only). If a GOP header does follow the sequence header 302, the bitstream is definitely a non-progressive refresh sequence. Otherwise, the bitstream may, or may not, be of progressive refresh. The picture layer 330 then follows.

[0083] The picture immediately following the sequence header (i.e., the first picture in the GOP, i.e., picture 324) can be either an I or P-picture. If it is an I-picture, the bitstream is a non-progressive refresh sequence. Otherwise, the bitstream is a progressive refresh sequence. The bit allocation procedures of the present invention for bitstreams with progressive refresh and non-progressive refresh are slightly different. In the following two subsections, we detail both.

[0084] The following notation will be used:

N,M - the actual values (i.e., reflecting the true GOP structure) of the GOP parameters, N and M;

N',M' - the initial assumed values of the GOP parameters, N and M;

N'',M'' - the (current) updated values of the GOP parameters, N and M;

N_I,N_P,N_B - the numbers of I, P and B pictures, respectively, in a GOP;

n_I,n_P,n_B - the number of I, P and B pictures, respectively, remaining in the current GOP;

R - the number of remaining bits for the current GOP;

bit_rate - the effective output bit rate in bits per second that may be variable (we assume that bit_rate is available to transcoder);

frame_rate - the *effective* frame_rate in frames per second, which may not be the same as the frame_rate embedded in the sequence header due to the presence of the repeat_first_field (we assume that frame_rate is available to transcoder);

div - Integer division with truncation of the result toward zero; and

⇐ - Assignment statement (from the right to the left).

(i) Progressive Refresh

[0085] For a progressive refresh sequence, there will be no GOP header, and the first picture immediately following the sequence header (in encoding order) has to be a P-picture. See sequences GOP(M=1) and GOP(M) in Table 1(b) and $GOP_0(M)$ in Table 2(b). Instead of waiting to correctly determine M, we proceed to process the first P-picture with any reasonable assumed M, say M', and a bit budget of $R = R_{GOP,M} = M' \cdot (bit\_rate/frame\_rate)$ bits in calculating the target bit rate (see eqn. 6). Note that with M', we have $N_P = 1$ and $N_B = M' - N_P$. For example, refer to FIG. 1(a) or 2(a), where the distance between P-pictures is M'=3, the number of P-pictures in a GOP is one, and the

number of B-pictures in a GOP is two.

[0086] The second picture in the GOP, in encoding order, can be either a B-picture (*GOP(M)*) or a P-picture (*GOP(M* = 1) or $GOP_0(M)$), as shown in Tables 1(b) and 2(b), respectively.

[0087] In either case, we can calculate the actual value of *M* from the temporal references of the first two pictures of a GOP in encoding order. Specifically, let *temp_ref_picture* 1 and *temp_ref_picture* 2 be the temporal references for the first and second pictures, respectively. The temporal reference may be embedded in the picture header. If the second picture is a B-picture, e.g., sequence (*GOP(M)*) in Table 1(b), the correct value of *M*, say *M"*, is:

$$M" = temp\_ref\_picture\ 1 - temp\_ref\_picture\ 2 + 1. \qquad (12)$$

[0088] For example, in FIG. 1(b), consider P_{10} and B_8 as first and second pictures, respectively, in an encoded video sequence. Then, temp_ref_picture1 is the temporal reference of P_{10} (e.g., "10"), temp_ref_picture2 is the temporal reference of B_8 (e.g., "8"), and M"=10-8+1=3. Note that the temporal references refer to the display order of the pictures.

[0089] Otherwise, if the second picture is a P-picture, e.g., sequence (*GOP(M* = 1) in Table 1(b), or sequence $GOP_0(M)$) in Table 2(b), the correct value of the distance between P-pictures is:

$$M" = temp\_ref\_picture\ 2 - temp\_ref\_picture\ 1. \qquad (13)$$

[0090] For example, in FIG. 2(b), consider P_0 and P_3 as first and second pictures, respectively, in an encoded video sequence. Then, temp_ref_picture1 is the temporal reference of P_0 (e.g., "0"), temp_ref_picture2 is the temporal reference of P_3 (e.g., "3"), and M"=3-0=3.

[0091] With a correct value of *M*, say *M"*, the remaining number of bits, *R* (see eqn. 7), is adjusted accordingly. Specifically, if the second picture is a B-picture, as with sequence (*GOP(M)*) in Table 1(b), R is updated as

$$R = R + (M" - M') \cdot \frac{bit\_rate}{frame\_rate}. \qquad (14)$$

[0092] As seen, for *M"*>*M'* (actual distance between P-pictures is greater than assumed distance), the first P-picture has been assigned relatively more bits and, therefore, the remaining pictures will be assigned relatively fewer bits. For *M"*<*M'* (actual distance between P-pictures is less than assumed distance), the first P-picture has been assigned relatively fewer bits and the remaining pictures will be assigned relatively more bits.

[0093] On the other hand, if the second picture is a P-picture, e.g., as with sequences (*GOP(M* = 1) in Table 1(b), or sequence $GOP_0(M)$) in Table 2(b), there are no B-pictures between the first two P-pictures in the encoded video sequence. Hence, the bits allocated for the assumed *M'*-1 B-pictures between two P pictures (in encoding order) should be subtracted from *R*, as

$$R = R - (M' - 1) \cdot \frac{bit\_rate}{frame\_rate}. \qquad (15)$$

[0094] We then set *M* = *M'*. If the transcoder can meet the target rate at each picture, at the end of the first *M* pictures, we will have *R* = 0. Hence, we will have a bit budget $R = R_{GOP,M} = M \cdot (bit\_rate/frame\_rate)$ bits for next *M* pictures. To address any subsequent change in *M*, we check the value of *M* after each P-picture.

[0095] FIG. 4(a) illustrates a coding method for a progressive refresh picture sequence in accordance with the present invention. The aforementioned bit allocation technique for progressive refresh sequences is summarized as follows.

[0096] At block 400, the pre-compressed digital video bitstream is received at a transcoder. The coded pictures are in an encoding order. At block 405, the bitstream is processed to detect a sequence header. At block 410, the picture type of the first picture following the sequence header is determined. If it is an I-picture, processing continues at block "A" in FIG. 4(b).

[0097] If the first picture following the sequence header is a P-picture, processing continues at block 415, where an assumed distance M' is set. M' is defined herein as a distance between the first picture following the sequence header, and the next P-picture in the display order. At block 420, an assumed bit budget R is determined based on M', the bit rate and the frame rate of the bitstream. At block 422, a target bit rate T is set in accordance with eqn. 6b. At block 425, the first picture is coded according to T. At block 430, if the next picture (e.g., the second picture) is a P-picture, the actual distance M" is determined at block 435 according to eqn. 13. At block 440, the bit budget R is adjusted using eqn. 15. At block 445, M is set to M".

[0098] At block 430, if the next picture in the sequence, e.g., the second picture, is a B-picture, the actual distance M" is determined at block 450 according to eqn. 12. At block 455, the bit budget R is adjusted using eqn. 14. At block 460, M is set to M". At block 462, the target bit rate T is set (e.g., using eqn. 6). Next, the remaining M-1 pictures (including the second picture) are coded according to T at block 465.

[0099] At block 470, the next picture is processed. This may be the first picture of the next series of M" pictures, for example. At block 475, processing returns, e.g., to block 415. However, note that, instead of blindly assuming the length and bit budgets at blocks 415 and 420, respectively, the values determined in the previous series of M" pictures may be used. Nevertheless, the actual distance M" may be verified, e.g., after every P-picture.

(ii) Non-Progressive Refresh

[0100] For non-progressive refresh, there may be an optional GOP header following the sequence header. Then, a picture layer follows. The first picture immediately following a sequence header, or a GOP header, has to be an I-picture. For an intelligent bit allocation, we need to know both N and M. We can, however, start transcoding of a pre-compressed bitstream with any reasonable pair of N and M, say N' and M', and with a bit budget of $R = R_{GOP,N} = N' \cdot (bit\_rate/frame\_rate)$ bits in determining the target rate for the first I-picture (see eqn. 6).

[0101] Note that with N' and M', we have $N_I = 1, N_P = N'/M' - N_I$ and $N_B = N' - N_I - N_P$. The actual value of M, say M'', can then be calculated from the temporal references of the first two pictures, similar to progressive refresh. If the second picture is an I- or P-picture, eqn. (13) is used in calculating the value of M. Otherwise, the second picture is a B-picture, and eqn. (12) is used. We then set $M = M''$.

[0102] If the second picture is also an I-picture ($GOP(N = M = 1)$ and $GOP_0(N = M)$), we actually have had the correct N value too, i.e., $N = M$, as shown in Table 1(b) and 2(b). The remaining number of bits, R (see eqn. 7), is adjusted accordingly, as:

$$R = R - (N' - 1) \cdot \frac{bit\_rate}{frame\_rate}. \qquad (16)$$

[0103] This is because there are no P- and B-pictures between the first and second I-pictures. Hence, the bits allocated to the assumed N-1 (P and B) pictures between two I pictures should be deducted from R.

[0104] If the second picture is not an I-picture, it will take somewhat longer to obtain the correct value of N, by about one GOP. The assumed N' may not be dividable by the correct M. Hence, we need to adjust N', and accordingly, n_P and nB. We adjust N to N'' with the correct M value, and n_P and n_B, as

$$N'' = (N'/M) \cdot M; \qquad (17)$$
$$n_p = N''/M - 1; \text{ and} \qquad (18a)$$
$$n_B = N''/M - n_p - 1. \qquad (18b)$$

[0105] We then adjust the remaining number of bits, R (see eqn. 7), with N'' as

$$R = R + (N'' - N') \cdot \frac{bit\_rate}{frame\_rate}. \qquad (19)$$

[0106] Furthermore, if the second picture is a P-picture, GOP(N,M=1) or GOP_0(N,M) as shown in Tables 1(b) and 2(b), the remaining number of bits, R, needs additional adjustment as

$$R = R - (M - 1) \cdot \frac{bit\_rate}{frame\_rate} \qquad (19b)$$

[0107] This is because there are M-1 fewer B-pictures than assumed before the first I-picture.

[0108] There can be only three scenarios, i.e.,

1. N''=N, that is, the current GOP length N'' is equal to the actual GOP length N;
2. N''<N, that is, the current GOP length N'' is smaller than the actual GOP length N; and
3. N''>N, that is, the current GOP length N'' is greater than the actual GOP length N.

[0109] Clearly, if $N = N''$, we will reach an I-picture after processing N'' pictures - the beginning of the second GOP. We have a correct N, so we maintain it.

[0110] If $N'' < N$, the GOP length is shorter than the actual GOP length. The assumed GOP length is at least M pictures shorter than the actual GOP length because of the nature of the GOP structure (see Tables 1(b) and 2(b)). Here, we have assumed that the value of M will not change within a GOP.

[0111] Note that M may be checked periodically within a GOP, such as after each I- and/or P-picture, if there is a chance that M may change within a GOP. Additionally, if M changes within the GOP, N and R should be adjusted accordingly.

[0112] We extend the GOP of N'' pictures by M additional pictures, i.e.,

$$N'' = N'' + M, \qquad (21)$$

and adjust R by adding $R_{GOP,M}$ additional bits for the M additional pictures, i.e.,

$$R = R + R_{GOP,M} = R + M \cdot \frac{bit\_rate}{frame\_rate}. \qquad (22)$$

[0113] If the extended GOP length is still shorter than the actual GOP length, we repeat the above procedure until reaching the end of the actual GOP.

[0114] If $N''>N$, the GOP length is longer than the actual GOP length. At the end of the actual GOP, we will have a correct GOP length, N, and we will use the correct length N for the rest of the GOPs. Since the assumed GOP length (N'') is longer than the actual GOP length (N), there will be some bits left over from the first actual GOP for coding pictures in the second actual GOP. Hence, the bits allocated for the second actual GOP need to be adjusted as:

$$R = 2R_{GOP,N'}-(R_{GOP,N''}-R)=2N\cdot\frac{bit\_rate}{frame\_rate}-(N''\cdot\frac{bit\_rate}{frame}-R) \qquad (23)$$

where $R_{GOP,N''}-R$ is the number of bits used for encoding the first N pictures.

[0115] FIG. 4(b) illustrates a coding method for a non-progressive refresh picture sequence in accordance with the present invention. The aforementioned bit allocation technique for non-progressive refresh sequences is summarized as follows.

[0116] At block "A" processing continues from block "A" of FIG. 4(a). At block 500, an assumed GOP length N' is provided, and at block 502, an assumed distance M' is provided. M' here is a distance between the first I-picture in a GOP and the next I- or P-picture in the display order. This distance may or may not be the same as the GOP length. Generally, compatible values of N' and M' should be selected, for example, N'>M' and N' is dividable by M'. At block 504, an assumed bit budget R is provided according to N', and at block 505, a target bit rate T is determined, e.g., using eqn. 6, for coding the pictures in the GOP.

[0117] At block 506, the first picture in the GOP is coded. At block 508, if the next picture, e.g., the second picture of the GOP, is an I-picture, processing continues at block 510, where the actual distance M" is determined using eqn. 13. At block 514, M is set to M", and at block 516 N is set to M since the distance M" and the GOP length are the same.

[0118] At block 518, the bit budget is adjusted according to eqn. 16, and at block 524 processing returns, e.g., to block 500. Note that, at blocks 500 and 502, N' and M' can be set to the values N and M, respectively, of the previous GOP. N and M can be verified for each GOP.

[0119] At block 508, if the next picture, e.g., the second picture in the GOP, is a P-picture, processing continues at block 530, where the actual distance M" is determined using eqn. 13. At block 532, M is set to M". At block 534, the GOP length is updated as N"=M[N'/M']. At block 536, the remaining number of B-pictures in the current GOP, n_B, is adjusted using eqn. 18. At block 538, the bit budget is adjusted using eqn. 19, and at block 540, the bit budget is adjusted again, using eqn. 19b, and processing continues at "B".

[0120] At block 508, if the next picture, e.g., the second picture in the GOP, is a B-picture, processing continues at block 542, where the actual distance M" is determined using eqn. 12. At block 544, M is set to M". At block 546, the GOP length is updated as N"=M[N'/M']. At block 548, the remaining number of B-pictures in the current GOP, n_B, is adjusted using eqn. 18. At block 550, the bit budget is adjusted using eqn. 19, and processing continues at "B".

[0121] FIG. 4(c) illustrates a continuation of the coding method of FIG. 4(b) in accordance with the present invention. Processing continues at "B" from FIG. 4(b). At block 552, the target bit rate T is set using eqn. 6, and at block 554, a counter is set to one. At block 556, the next picture in the GOP is coded. At block 558, counter is incremented. "Counter" essentially tracks the position of the current picture in the GOP. At block 560, a determination is made as to whether the end of the current GOP has been reached. If so, a determination is made at box 562 as to whether counter=N". If so, this indicates that N" is the same as N, the actual GOP length. Essentially, the assumed GOP length N" turned out to be correct. Processing continues at block 566, where the next GOP is coded.

[0122] If the condition at block 562 is false, this indicates that N">N (e.g., the assumed GOP length N" is too large), and the bit rate is adjusted at block 564 using eqn. 23. Processing continues at block 566, where the next GOP is coded.

[0123] If the end of the current GOP has not been reached yet at block 560, a determination is made at block 570 as to whether counter=N". If not, process continues at block 556. If counter=N" at block 570, this indicates that N"<N (e.g., the assumed GOP length N" is too small). At block 572, the assumed GOP length N" is incremented by M. at block 574, the bit budget is adjusted using eqn. 22, at block 576, the next M pictures are coded, and at block 578 a determination is made as to whether the end of the current GOP has been reached. If not, processing continues at block 572, where the assume GOP length N" is again incremented by M.

[0124] When the end of the current GOP is reached, at block 578, processing continues at block 566, where the next GOP is coded.

[0125] The coding method of FIG. 4(c) can be understood further with reference to FIG. 5.

[0126] FIG. 5 illustrates an estimated GOP length N" in accordance with the present invention. A sequence of pictures is shown generally, in display order, at 500. A first GOP, GOP(N,M) 520, begins at an I-picture 502, and the next GOP 580 begins at an I-picture 510. A subsequent GOP begins at an I-picture 515.

[0127] The estimated GOP length N" 530 is equal to the actual GOP length N. However, the estimated GOP length N" 540 is less than the actual GOP length N by two M distances 542 and 544. In particular, the estimated GOP length N" 540 extends from the I-picture 502 to a P-picture 504, the M-distance 542 extends from the P-picture 504 to a P-picture 506, and the M-distance 544 extends from the P-picture 506 to the I-picture 510. The estimated GOP length N" 550 extends from the I-picture 502 to a P-picture 512, and is therefore greater than the actual GOP length N.

Generally, N" should be selected initially as a multiple of the estimated distance between I and/or P-pictures.

[0128] A C-language pseudo code for the bit-allocation procedure, both with and without progressive refresh, is shown below. The same algorithm is represented by the flowchart in FIGs 7(a)-7(d). A single bitstream can switch arbitrarily between the progressive refresh and I-picture (non-progressive) refresh modes.

```
/*
 * The following is a brief description of the variables:
 *
 * N    Estimated open GOP length (pictures)
 * M    Number of consecutive B-pictures plus 1
 * N'   Initial guess for the GOP length in pictures
 * M'   Initial guess for the number of consecutive B-pictures plus 1
 * N"   Updated estimate of the open GOP length
 * M"   Updated estimate of number of consecutive B-pictures plus 1
 * R    Number of bits remaining in this GOP
 * n    Number of pictures remaining in this GOP
 * ni   Number of I-pictures remaining in this GOP
 * np   Number of P-pictures remaining in this GOP
 * nb   Number of B-pictures remaining in this GOP
 * Xi   Complexity of an I-picture
 * Xp   Complexity of a P-picture
 * Xb   Complexity of a B-picture
 * Ki   I-picture complexity divider
 * Kp   P-picture complexity divider
 * Kb   B-picture complexity divider
 * X    Complexity temporary variable
 * T    Target number of bits for the current picture type
 *
 * pict_type     Current picture type
 * prev_pict_type     Previous picture type in coding order
 * temp_ref           Current picture temporal reference
```

```
* prev_temp_ref      Previous picture temporal reference (coding order)
* prog_refresh       Non-zero if progressive refresh is in progress
* GOP_flag           Non-zero if GOP header preceded the current picture
* seq_hdr_flag       Non-zero if sequence header preceded current picture
* bit_rate()    Current average output bitrate
* frame_rate()       Average display frame rate using repeat_first_field
*
* Notes:
*
*  1) The "div" operation denotes computer integer division with
*       truncation toward 0.
*
*  2) A GOP begins with a I-picture in coding order and extends to
*       the next I-picture (exclusive) regardless of whether the GOP
*       syntax is present in the bitstream.  When progressive refresh
*       is used a GOP consist of a P-picture plus the following run of
*       consecutive B-pictures.
*/

FindSequenceHeader();
N = N';          // (N' % M') == 0; N is estimated open GOP length
M = M';
R = 0;
n = 0;
prev_pict_type = B_PICTURE;
for (;;) {
        /*
```

```
        * get_picture_header() parses the video sequence and the optional
        * GOP layer, and returns the picture type and temporal reference
        * of the next coded picture.
        */
       pict_type = get_pict_header(&temp_ref, &GOP_flag, &seq_hdr_flag);
       if (seq_hdr_flag)
              prog_refresh = (pict_type != I_PICTURE);
       switch (pict_type) {

       case I_PICTURE:
              if (prev_pict_type != B_PICTURE) {
                     /*
                      * GOP(N=M=1), GOP(N,M=1) or GOP0(N=M).  Recalculate
M,
                      * start new GOP and correct R for missing B pictures.
                      */
                     M" = GOP_flag ? (temp_ref + 1) :
                            ((temp_ref – prev_temp_ref + 1024) & 0x3FF);
                     R -= n * bit_rate() / frame_rate();
                     M = M";
                     N = (prev_pict_type == I_PICTURE) ? M : (N - n);
                     n = 0;
              }
              if (prog_refresh)              // End of progressive refresh
                     N = N';
              prog_refresh = 0;
              if (n <= 0) {    // GOP length is equal to expected length
```

```
                    R += N * bit_rate() / frame_rate();
                } else {                    // GOP length is shorter than expected
                    N -= n;
                    N = (N div M) * M;
                    R += (N – n) * bit_rate() / frame_rate();
                }
                n = N;
                ni = 1;
                np = N/M - 1;
                nb = N - ni - np;
                X = Xi / Ki;
                prev_temp_ref = temp_ref;
                break;

        case P_PICTURE:
                if (prev_pict_type != B_PICTURE) {
                    /*
                     * GOP(M=1) or GOP0(M)                    [progressive
refresh]
                     * GOP(N,M=1) or GOP0(N,M)                [non-progressive
refresh]
                     */
                    M" = (temp_ref - prev_temp_ref + 1024) & 0x3FF;
                    R  -= (M - 1) * bit_rate() / frame_rate();
                    nb -= M - 1;
                    n  -= M - 1;
                    if ((M" != M) && !prog_refresh) {
```

```
                        N" = (N div M") * M";          // open, not actual, GOP
length
                        n +=  (N" - N) - (M" - M);
                        R += ((N" - N) - (M" - M)) * bit_rate() / frame_rate();
                        np = n div M";
                        nb = n - np;
                        N  = N";
                    }
                    M = M";
                }
                if (n <= 0) {
                    /*
                     * End of GOP in progressive refresh mode or GOP is
                     * longer than expected in non-progressive refresh
mode.
                     */
                    N += M;
                    R += M * bit_rate() / frame_rate();
                    n = M;
                    ni = 0;
                    np = 1;
                    nb = M - 1;
                }
                X = Xp / Kp;
                prev_temp_ref = temp_ref;
                break;
```

```
        case B_PICTURE:
                if (prev_pict_type != B_PICTURE) {
                        /*
                         * GOP(N=M), GOP(N,M), GOP0(N,M)  [non-progressive
refresh]
                         * GOP(M)
[progressive refresh]
                         */
                        M" = (prev_temp_ref - temp_ref + 1025) & 0x3FF;
                        if (M" != M) {
                                if (prog_refresh) {
                                        nb += M" - M;
                                        R += (M" - M) * bit_rate() / frame_rate();
                                } else {
                                        N" = (N div M") * M";
                                        n +=  N" - N;
                                        R += (N" - N) * bit_rate() / frame_rate();
                                        np = (n + 1) div M" - 1;
                                        nb = n - np;
                                        N = N";
                                }
                                M = M";
                        }
                }
                X = Xb / Kb;
                break;
        }
```

```
        /*
         * Calculate the target bit rate
         */
        T = X * R / (ni*Xi/Ki + np*Xp/Kp + nb*Xb/Kb);
        switch (pict_type) {            // Decrement pic count for each picture type
                case I_PICTURE: ni--; break;
                case P_PICTURE: np--; break;
                case B_PICTURE: nb--; break;
        }
        n = n - 1;                      // Number of pictures left in GOP
        prev_pict_type = pict_type;
        /*
         * transcode_picture() does the actual transcoding which maintains
         * in original picture type (I, P or B). This function returns
         * the actual size of the transcoded picture in bits. This func-
         * tion updates the complexity (Xi, Xp, Xb) for each picture type.
         */
        R -= transcode_picture(pict_type, T);
}
```

[0129] This pseudo-code is expressed in a flowchart 700 in FIGs 7(a)-7(d). Note that ni, np and nb are the same as n_I, n_P and n_B, respectively, as used elsewhere.

[0130] Box 702 denotes searching for the video sequence header in order to obtain parameters necessary to begin transcoding. Initialization for the loop over transcoded pictures occurs here.

[0131] Box 704 is the call to "get_pict_header()" which consumes the picture header and all higher level syntax. The picture type (I, P or B picture) is returned. "GOP_flag" is true (non-zero) if a GOP header was encountered; "seq_hdr_flag" is true (non-zero) if a video sequence header occurred before the current coded picture and after the previous coded picture.

[0132] Box 706 tests "seq_hdr_flag". If true, "prog_refresh" is set to 1 (true) if the current picture is not an I-picture in box 708

[0133] Box 710 is the switch statement based on picture types. Although not shown in the code or flowchart, an MPEG-2 D-picture is treated like an I-picture in this algorithm.

[0134] If the previous picture type was not a B-picture (box 712), then a new value of M is determined from the temporal references. If the GOP is being shortened, the excess bits due to the expected GOP length are deleted from "R" and the new values of "M", "N", "n" are set (box 714). "R" is the target number of bits (not the actual number) remaining to code the rest of the GOP. If the actual GOP size is different, the difference in bits (positive or negative) remains in "R" and affects the target size for the next GOP. This forms a feedback loop to drive the average GOP size to "N*bit_rate() / frame_rate()".

[0135] The function "bit_rate()" returns the current target transcoder output bitrate which may change due to channel capacity or statistical multiplexing considerations. The function "frame_rate()" is the average coded frame rate, which is the frame rate specified in the video sequence header modified by the effects of the repeat_first_field flag in the various coded pictures. The quantity "bit_rate() / frame_rate()" denotes the desired current average coded picture size (in bits) to be output by the transcoder. It is an average because it does not consider the different MPEG picture types. Thus the target number of bits for a GOP, "N * bit_rate() / frame_rate()",

is N (the GOP length in pictures) times the average picture size.

[0136] If the transcoder is operating in progressive refresh mode, this mode is disabled because the current picture is an I-picture and N is set to the initial guess for GOP length (boxes 716 and 718).

[0137] If the condition in 720 is true, the correct value of "N" has been used. In this case, R is updated with the number of bits needed for a new GOP of length N (box 722). If the 720 condition is false, the our previous estimate of "N" is wrong (the GOP is shorter than expected) and N and R are updated accordingly (box 724).

[0138] Box 726 resets the picture type down counters ("ni", "np", "nb") according to the GOP type, the number of picture to go in the GOP (n), calculates the normalized I-picture complexity (X) and updates the previous picture type.

[0139] Box 728 is the start of the P-picture case in the switch statement. If the previous picture was not a B-picture (box 730), then a new value of M (M") is calculated from the temporal references. The difference between the two 10-bit temporal reference number will be negative a wrap around occurred; this is why 1024 is added. When no wrap around occurs, the 0x3FF, which is the hexadecimal representation of the binary number 1111111111, removes the 2^{10}à bit leaving the positive difference; otherwise the addition of 1024 make the negative difference positive and this result is not modified by the bit-wise AND operation. Since the previous picture was not a B-picture, potential corrections to "n" (number of pictures left in the GOP), "nb" (number of B-picture left in the GOP) and "R" (number of bits allocated for the remaining picture in the GOP) may be required.

[0140] If M has changed and the transcoder is not in progressive refresh mode (box 732), then corrections to "N", "n", "np", "nb" and "R" are made in box 734.

[0141] Box 736 updates M (previously calculated in M").

[0142] If the end of GOP has been reached, (n less than or equal to zero - box 738), then the GOP is actually longer than expected. In this case, the values of "N", "R", "n", "ni", "np" and "nb" are updated to reflect an increase GOP length of M pictures (one P-picture and M-1 B-pictures) in box 740.

[0143] Box 742 calculates the normalized P-picture complexity, "X", and resets "prev_temp_ref" to be the current temporal reference.

[0144] For B-pictures, where the previous picture was not a B-picture (conditional in box 744), the new value of M (in M") is computed. (see box 746) from the difference in temporal reference. This calculation provide for the possibility that the 10-bit temporal reference may have wrapped around (hence the +1024 and the mask to 10 bits).

[0145] If the value of M has changed (condition tested in box 748), then the state variables must be updated depending on "prog_refersh" (box 750). In the progressive refresh case, "nb" and "R" need to be modified (box 752). If I-picture refresh is in progress, "N", "np", "nb", "n" and "R" all need to change (box 754).

[0146] A new normalized complexity for B-picture is computed in box 758.

[0147] Following the switch statement, the target size, "T", for the next picture is calculated in box 760. This number is the intended size of the current picture when transcoded to the new bit rate. The actual size, returned by "transcode_picture()" is used to decrement R (see box 770). Following the calculation of "T", the number of picture remaining in the GOP, "n", and the number of pictures of the current type ("ni", "np" or "nb") must be decremented (see boxes 762 to 768).

3. DIFFERENCE IN TARGET RATES

[0148] The target bit rates for a picture determined with different sets of GOP parameters can be different (see equations 5 and 6). We now analyze the differences in target rates with and without a correct starting GOP structure. For sake of simplification in analysis, we assume that:

1. The transcoder can meet the target rate at each picture. This implies that the remaining number of bits at the end of a GOP is equal to zero, and that a GOP of N pictures is assigned exactly $R_{GOP,N} = N \cdot (bit\_rate/frame\_rate)$ bits.

2. Within a GOP, all the pictures of the same type have the same complexity measure.

3. We have a correct M value. In fact, we can have a correct M right after reading the header of the second picture.

[0149] Note that the above assumptions are only for simplifying the analysis, and may not be 100% true for a real transcoder. For example, a transcoder may not 100% meet the target rate at a frame, and therefore at the end of a GOP, the remaining number of bits R may not be zero bits.

[0150] It has been shown that with conditions 1 and 2, equations (5) and (6) will be identical. That is, all the pictures of the same type in a GOP are assigned the same number of bits. We use eqn. (5) in our analysis of differences in target rates. Assume that a pre-compressed bitstream is coded with a set of GOP parameters as follows:

$$N,M \rightarrow N_I, N_P, N_B$$

[0151] For comparison, we start transcoding of the pre-compressed bitstream with two different GOP structures, one with a correct GOP structure (I) and the second with an incorrect GOP structure (II) as follows:

Starting GOP(I): $N,M \rightarrow N_I,N_P,N_B$
Starting GOP(II): $N',M' \rightarrow N_I,N_P,N_B$

[0152] Let T_n and T'_n be the target rates for frame n with GOP(I) and GOP(II), respectively. The difference between these two target rates is:

$$\Delta=T_n-T'_n=(1-\frac{T'_n}{T_n})\cdot T_n \qquad (24a)$$

$$=(1-\frac{\frac{X_{n,t}}{K_{n,t}}}{N'_I\frac{X_I}{K_I}+N'_P\frac{X_P}{K_P}+N'_B\frac{X_B}{K_B}}R_{GOP,N'}\div\frac{\frac{X_{n,t}}{K_{n,t}}}{N_I\frac{X_I}{K_I}+N_P\frac{X_P}{K_P}+N_B\frac{X_B}{K_B}}R_{GOP,N})\cdot T_n$$

(24b)

$$=(1-\frac{N_I\frac{X_I}{K_I}+N_P\frac{X_P}{K_P}+N_B\frac{X_B}{K_B}}{N'_I\frac{X_I}{K_I}+N'_P\frac{X_P}{K_P}+N'_B\frac{X_B}{K_B}}\cdot\frac{N'}{N})\cdot T_n \qquad (24c)$$

$$=\frac{(NN'_I-N'N_I)\frac{X_I}{K_I}+(NN'_P-N'N_P)\frac{X_P}{K_P}+(NN'_B-N'N_B)\frac{X_B}{K_B}}{(N'_I\frac{X_I}{K_I}+N'_P\frac{X_P}{K_P}+N'_B\frac{X_B}{K_B})}\cdot\frac{1}{N}\cdot T_n .$$

(24d)

[0153] To better understand this difference in target bit rate, let us examine two numerical examples. In the first example, we assume the starting GOP (II) is shorter than the actual GOP, i.e.,

$$N=15\ M=3\rightarrow N_I=1\ N_P=4\ N_B=10$$
$$N'=12\ M'=3\rightarrow N'_I=1\ N'_p=3\ N'_B=8$$

[0154] If we set $K_I=K_P=1$ and $K_B=1.4$, and use the complexity measures as initialized in MPEG TM5, i.e.,

$$X_I=160\cdot bit\_rate/frame\_rate$$
$$X_P=60\cdot bit\_rate/frame\_rate \qquad (25)$$
$$X_B=40\cdot bit\_rate/frame\_rate$$

we have the difference between the two target rates as follows:

$$\Delta=T_n-T'_n=\frac{(15-12)\ x_I\ +\ (15\cdot3-12\cdot4)x_P\ +\ (15\cdot8-12\cdot10)x_B/1.4}{(x_I\ +\ 3x_P\ +\ 8x_B/1.4)}\cdot\frac{1}{15}\cdot T_n=3.5\%\cdot T \qquad (26)$$

Here, the target rate with *a priori* knowledge of the GOP structure is 3.5% higher than the rate without *a priori* knowledge of the GOP structure using the rate control system of the present invention. A shorter GOP results in smaller target rates for all three types of pictures.

[0155] In the second numerical example, the starting GOP (II) has a length longer than the actual one, i.e.,

$$N=15\ M=3\rightarrow N_I=1\ N_P=4\ N_B=10$$
$$N'=18\ M'=3\rightarrow N'_I=1\ N'_p=5\ N'_B=12$$

[0156] The difference in target rates is

$$\Delta = T_n - T'_n = \frac{(15-18)\,x_I + (15\cdot 5-18\cdot 4)x_P + (15\cdot 12-18\cdot 10)x_B/1.4}{(x_I + 5x_P + 12x_B/1.4)} \cdot \frac{1}{15} \cdot T_n = -1.7\% \quad (27)$$

[0157] Here, the target rate with *a priori* knowledge of the GOP structure is 1.7% lower than the rate without *a priori* knowledge of the GOP structure using the rate control system of the present invention. A longer GOP results in the assignment of more bits to all the pictures of three types.

[0158] The differences in target bit rates will remain until the end of either the starting GOP or the actual GOP, depending on whether the starting GOP is shorter or longer than the actual GOP length.

[0159] If the starting GOP is shorter than the actual GOP (i.e. $N' < N$), at the end of the starting GOP, we will extend the GOP of N' pictures by M pictures with additional bits of $R_{GOP}M$. The remaining number of bits is therefore equal to

$$R = R + R_{GOP,M} = M \cdot \frac{bit\_rate}{frame\_rate} \quad (28)$$

[0160] Note that we have assumed $R = 0$ at the end of the starting GOP of N' pictures. The additional M pictures actually form another GOP of $N = M$ pictures with $N'_I = 0$, $N'_P = 1$ and $N'_B = M-1$, and a bit budget of $M \cdot (bit\_rate/frame\_rate)$ bits. The differences in target rates for the M additional pictures of either P or B type (see eqn. 24) can be simplified as:

$$\Delta = T_n - T'_n = \frac{(-MN_I)\frac{X_I}{K_I} + (N - MN_P)\frac{X_P}{K_P} + (NN'_B - MN_B)\frac{X_B}{K_B}}{(\frac{X_P}{K_P} + N'_B\frac{X_B}{K_B})} \cdot \frac{1}{N} \cdot T_n \quad (29)$$

[0161] For the first numerical example, the difference in target bit rates is:

$$\Delta = T_n - T'_n = \frac{(-3)\,x_I + (15-3\cdot 4)x_P + (15\cdot 2-3\cdot 10)x_B/1..4}{(x_P + 2x_B/1.4)} \cdot \frac{1}{15} \cdot T_n = 17.0\% \cdot T_n \quad (30)$$

[0162] Here, the target rate with *a priori* knowledge of the GOP structure is 17.0% higher than the rate without *a priori* knowledge of the GOP structure using the rate control system of the present invention. Relatively more bits are assigned to the pictures of the extended GOP since relatively fewer bits were assigned to the previous pictures. If the extended GOP of M pictures still cannot end the actual GOP of N pictures, we repeat the above procedure until reaching the end of the actual GOP. The differences in target rates for the pictures in the extended GOPs of M pictures will remain the same because all the extended GOPs of M pictures have the same pattern of picture organization and the same bit budget.

[0163] At the end of the first actual GOP of N pictures and also at the end of the last extended GOP of M pictures, we have complete information about the actual GOP structure with the remaining number of bits $R = 0$. Hence, from the second GOP, there will be no difference in target rates. Two target rates with and without *a priori* knowledge of GOP structure using the rate control system of the present invention will be the same from the second actual GOP.

[0164] On the other hand, if the assumed starting GOP length is longer than the actual GOP length (i.e. $N' > N$), at the end of the actual GOP, we have the complete information about the actual GOP structure, but the remaining number of bits $R \neq 0$. That is, there will be some left over bits to allocate over the second (actual) GOP. Hence, instead of simply giving $R_{GOP}N$ bits to the second (actual) GOP, we have to adjust R as $R = 2R_{GOP,N} - (R_{GOP,N} - R)$ (see eqn. 23). The target bit rates with correct and incorrect starting GOP structures will still not be the same in the second (actual) GOP. The difference in target rates in the second (actual) GOP is

$$\Delta = T_n - T'_n \quad (31a)$$

$$= \frac{\frac{X_{n,t}}{K_{n,t}}}{N_I\frac{X_I}{K_I}X_I + N_P\frac{X_P}{K_P} + N_B\frac{X_B}{K_B}}(R_{GOP,N} - (2R_{GOP,N} - (R_{GOP,N'} - R))) \quad (31b)$$

$$= \frac{\frac{X_{n,t}}{K_{n,t}}}{N_I \frac{X_I}{K_I} X_I + N_P \frac{X_P}{K_P} + N_B \frac{X_B}{K_B}} (-R_{GOP,N} + (R_{GOP,N'} - R)) \qquad (31c)$$

[0165] Here, R_{GOP} N-R is the number of bits assigned for the first N pictures, which is equal to:

$$R_{GOP,N'} - R = \frac{N_I \frac{X_I}{K_I} + N_P \frac{X_P}{K_P} + N_B \frac{X_B}{K_B}}{N'_I \frac{X_I}{K_I} + N'_P \frac{X_P}{K_P} + N'_B \frac{X_B}{K_B}} R_{GOP,N'} . \qquad (32)$$

[0166] Hence, the difference in target rate becomes:

$$\Delta = \frac{\frac{X_{n,t}}{K_{n,t}}}{N_I \frac{X_I}{K_I} + N_P \frac{X_P}{K_P} + N_B \frac{X_B}{K_B}} (-R_{GOP,N} + \frac{N_I \frac{X_I}{K_I} + N_P \frac{X_P}{K_P} + N_B \frac{X_B}{K_B}}{N'_I \frac{X_I}{K_I} + N'_P \frac{X_P}{K_P} + N'_B \frac{X_B}{K_B}} R_{GOP,N'}) \qquad (33a)$$

$$= -\frac{(NN'_I - N'N_I)\frac{X_I}{K_I} + (NN'_P - N'N_P)\frac{X_P}{K_P} + (NN'_B - N'N_B)\frac{X_B}{K_B}}{(N'_I \frac{X_I}{K_I} + N'_P \frac{X_P}{K_P} + N'_B \frac{X_B}{K_B})} \cdot \frac{1}{N} \cdot T_n \qquad (33b)$$

[0167] By comparing equations (24) and (33), we should see the differences in target rates in the first and second actual GOP have the same absolute values, but different signs. In fact, a longer starting GOP uses more bits for the first actual GOP and therefore leaves less bits for the second actual GOP. For the second numerical example, the differences in target rates for pictures in the second actual GOP are:

$$\Delta = T_n - T'_n = 1.7\% \cdot T_n \qquad (34)$$

[0168] At the end of the second actual GOP, the remaining number of bits R = 0. The third actual GOP will be assigned R_{GOP} N bits. From the third actual GOP, there will be no difference in target rates.

[0169] Target rates with respect to frame number for the two numerical examples were analyzed. The target rate with a correct starting GOP(N = 15, M = 3) was the benchmark for comparison. Comparative data was obtained for a starting GOP(N = 12, M = 3), and for a starting GOP(N = 18, M = 3). It was verified that all the pictures of the same type are assigned the same number of bits. Moreover, if the starting GOP is shorter than the actual GOP, fewer bits are assigned to the pictures before the end of the starting GOP, and more bits are assigned to the pictures of an extended GOP of M pictures. There is no difference in target rates starting from the second actual GOP.

[0170] On the other hand, if the starting GOP is longer than the actual one, slightly more bits are assigned to the pictures in the first actual GOP and fewer bits for the pictures in the second actual GOP. From the third actual GOP, the target rates are the same.

4. SIMULATION RESULTS

[0171] Simulations were conducted in evaluating the rate control system of the present invention. Test sequences were first coded at a bit rate of 15 Mbits/sec. with different sets of GOP parameters, as shown in Table I. The compressed bitstreams were then transcoded to a new rate of 3 Mbits/s with a starting GOP of N = 15 and M = 3.

	GOP1	GOP2	GOP3
N	30	12	8

	GOP1	GOP2	GOP3
M	3	2	1

[0172] Target rates were analyzed for a correct starting set of GOP parameters, and for a starting GOP of $N = 15$ and $M = 3$. Note that, in practice, a transcoder (or encoder) may not always meet the target rates at each frame. Hence, the remaining number of bits at the end of each GOP, R, may not be zero. The non-zero R will be passed to the next GOP (see eqn. 8). This implies that GOPs may be given slightly different number of bits. It was seen that a longer starting GOP assigns more bits for the pictures in the first GOP, as compared with a correct starting GOP. Hence, fewer bits are left for the pictures in the second GOP. The target rates with correct and incorrect GOP starts converging at the third GOP. On the other hand, a shorter starting GOP results in smaller target rates for the first few pictures and then more bits for the rest in the first GOP. Target rates become close from the second GOP.

[0173] To prove the effectiveness of the bit rate allocation system under a variety of test conditions, a sequence was coded with various GOP lengths, e.g., $N = 15, 12, 12, 18, 18, 15, 15, 9, 9, 30$. Target rates were analyzed at each GOP for the correct GOP length, and by using the previous GOP structure length as the starting GOP length for each GOP. The results were close with this stress test.

[0174] FIG. 6 illustrates a rate control system in accordance with the present invention. Pre-compressed bitstreams, e.g., program 1 and program 2, are provided to respective transcoders 620 and 630. One or several bitstreams and transcoders may be present. The bitstreams may be received real-time at the transcoders from a remote source, such as via a satellite broadcast, or may be provided from a local storage medium, for example.

[0175] The transcoders 620, 630 partially decompress the respective bitstreams, and encode the partially decompressed data at a different data rate, typically by using a different quantization parameter, according to a target bit rate signal provided by a rate control processor 605.

[0176] The rate control processor 605 receives information from the transcoders, including the picture type of the current picture, and the quantization parameter used to encode the current picture. This information is processed as described herein to set the target bit rates. The rate control processor 605 may include known processing circuitry, such as a central processing unit (CPU) 612, and a memory 614. The memory 614 may be a non-volatile memory that stores data, for example, for providing the initial distances M' and N' discussed previously. The rate control processor includes computing circuitry for executing the C-language like pseudo-code discussed previously.

[0177] A user interface 608 may communicate with the rate control processor 605 to set the default distances M' and N' or other relevant parameters in the rate control process.

[0178] The transcoded data from the transcoders 620, 630 is provided to a multiplexer (MUX) 660, and then to an encoder buffer 670. A buffer fullness signal is provided from the encoder buffer 670 to the rate control processor 605. The encoded data is finally transmitted over a channel at its new transcoded data rate.

5. CONCLUSION

[0179] A novel rate control system suitable for use with a digital video transcoder, such as one conforming to the MPEG standard, has been presented. The proposed rate control causes no processing delay and requires no extra memory. It can start with any reasonable set of GOP parameters, and then gradually corrects them when necessary as successive pictures are coded. Hence, it is able to address changes in the GOP structure of pre-compressed bitstreams, for example, when switching channels, inserting commercials, and the like. It has been shown that the target rates with correct and incorrect starting GOP parameters will converge within one or two GOPs. The differences in target rates are within a fairly small margin. Moreover, similar target rates correspond to a similar PSNR or picture quality.

[0180] Although the invention has been described in connection with various specific embodiments, those skilled in the art will appreciate that numerous adaptations and modifications may be made thereto without departing from the spirit and scope of the invention as set forth in the claims.

Claims

1. A method for allocating bits for coding pictures in a bitstream received at a digital video transcoder, comprising the steps of:

   (a) providing an assumed distance M' between a first picture of said bitstream and the next closest subsequent P-picture of said bitstream in a display order of said bitstream;

   (b) providing an assumed bit budget as a function of said assumed distance M';

   (c) coding said first picture in accordance with said assumed bit budget;

(d) determining a picture type of a second picture that immediately follows said first picture in said bitstream in an encoding order of said bit stream;

(e) adjusting said assumed bit budget according to said picture type of said second picture; and

(f) allocating bits for coding said second picture in accordance with said adjusted bit budget.

2. The method of claim 1, wherein:

said assumed bit budget is proportional to said assumed distance and a bit rate of said bitstream, and inversely proportional to a frame rate of said bitstream.

3. The method of claim 1 or 2, wherein:

said pictures in said bitstream form a progressive refresh sequence.

4. The method of one of claims 1 to 3, wherein said first picture is a P-picture and said second picture is a B-picture, comprising the further step of:

determining an actual distance M" between said first picture and said next closest subsequent P-picture, in said display order, according to said picture type of said second picture.

5. The method of claim 4, wherein:

said actual distance M" is determined according to a difference between a temporal reference of said second picture, in said display order, and a temporal reference of said first picture, in said display order, plus one picture.

6. The method of claim 4 or 5, comprising the further step of:

allocating bits for coding the remaining M"-1 pictures following said first picture in accordance with said adjusted bit budget.

7. The method of one of claims 4 to 6, wherein:

said adjusting step (e) comprises the step of adjusting said assumed bit budget by +(M"-M')*"bit rate"/"frame rate", where "bit rate" is a bit rate of said bitstream, and "frame rate" is a frame rate of said bitstream.

8. The method of claim 1, wherein said first and second pictures are P-pictures, comprising the further step of:

determining an actual distance M" between said first picture and said next closest subsequent P-picture, in said display order, according to said picture type of said second picture.

9. The method of claim 8, wherein:

said actual distance M" is determined according to a difference between a temporal reference of said second picture, in said display order, and a temporal reference of said first picture, in said display order.

10. The method of claim 8 or 9, wherein said adjusting step (e) comprises the step of:

adjusting said assumed bit budget by -(M'-1)*"bit rate"/"frame rate", where "bit rate" is a bit rate of said bitstream, and "frame rate" is a frame rate of said bitstream.

11. The method of one of claims 8 to 10, comprising the further step of:

allocating bits for coding the remaining M"-1 pictures following said first picture in accordance with said adjusted bit budget.

12. The method of one of claims 1 to 5 or 8 to 11, comprising the further step of:

allocating bits for coding a series of M" pictures following an initial M" picture that includes said first and second pictures in said bitstream in accordance with said adjusted bit budget.

13. A method for allocating bits for coding a sequence of pictures in a bitstream received at a digital video transcoder, comprising the steps of:

(a) providing an assumed length N' of a particular group of pictures (GOP) of said bitstream;

(b) providing an assumed bit budget as a function of said assumed length N';

(c) allocating bits for coding a first picture of said particular GOP in accordance with said assumed bit budget;

(d) determining a picture type of a second picture that immediately follows said first picture;

(e) adjusting said assumed bit budget according to said picture type of said second picture; and

(f) allocating bits for coding said second picture in accordance with said adjusted bit budget.

14. The method of claim 13, wherein:

said assumed bit budget is proportional to said assumed length N' and a bit rate of said bitstream, and inversely proportional to a frame rate of said bitstream.

15. The method of claim 13 or 14, wherein said first and second pictures are I-pictures, comprising the further steps of:

determining an actual distance M" between said first picture and the next closest subsequent I-picture of said bitstream according to the picture type of said second picture; and

adjusting said assumed bit budget in accordance with said actual distance M".

16. The method of claim 15, wherein:

said actual distance M" is determined according to a difference between a temporal reference of said second picture, in a display order of said bitstream, and a temporal reference of said first picture, in said display order.

17. The method of one of claims 13 to 16, wherein said adjusting step (e) comprises the step of:

adjusting said assumed bit budget by -(N'-1)*"bit rate"/"frame rate", where "bit rate" is a bit rate of said bitstream, and "frame rate" is a frame rate of said bitstream.

18. The method of one of claims 13 to 17, comprising the further step of:

allocating bits for coding the remaining pictures in said particular GOP following said second picture in accordance with said adjusted bit budget.

19. The method of one of claims 13 to 18, comprising the further step of:

providing an assumed distance M' between said first picture and the next closest subsequent P-picture of said bitstream, in a display order of said bitstream;
wherein said bits are allocated for coding said first picture in said step (c) in accordance with said assumed distance M'.

20. The method of claim 19, wherein said first picture is an I-picture and said second picture is a P-picture, comprising the further steps of:

determining an actual distance M" between said first picture and the next closest subsequent P-picture, in said display order, according to said picture type of said second picture; and

adjusting said assumed length N' of said GOP in accordance with said actual distance M" to provide an adjusted assumed length N"; wherein:

said adjusting step (e) is responsive to said adjusted assumed length N'.

21. The method of claim 20, wherein:

said assumed length N' of said particular GOP is adjusted in said adjusting step thereof by a factor, M(N'/M'), to provide said adjusted assumed length N", where M=M".

22. The method of claim 20 or 21, wherein said adjusting step (e) comprises the step of:

adjusting said assumed bit budget by a factor(N"-N')*"bit rate"/"frame rate", where "bit rate" is a bit rate of said bitstream, and "frame rate" is a frame rate of said bitstream.

23. The method of claim 20 or 21, wherein said adjusting step (e) comprises the further step of:

adjusting said assumed bit budget by a factor, (M"-1')*"bit rate"/"frame rate", where "bit rate" is a bit rate of said bitstream, and "frame rate" is a frame rate of said bitstream.

24. The method of one of claims 20 to 23, comprising the further step of:

allocating bits for coding the remaining pictures in said particular GOP following said second picture in accordance with said adjusted bit budget.

25. The method of one of claims 20 to 24, comprising the further steps of:

determining if N">N, an actual length of said particular GOP, after coding N" pictures in said particular GOP; and, if N">N:

further adjusting said adjusted bit budget according to N and N"; and

allocating bits for coding pictures in said bitstream following the N" coded pictures according to said further adjusted bit budget.

26. The method of one of claims 20 to 24, comprising the further steps of:

determining if N"<N, an actual length of said particular GOP, after coding N" pictures in said particular GOP; and, if N"<N:

further adjusting said adjusted bit budget according to M=M"; and

allocating bits for coding pictures in said bitstream following the N" coded pictures according to said further adjusted bit budget.

27. The method of claim 13, wherein said first picture is an I-picture and said second picture is a B-picture, comprising the further step of:

providing an assumed distance M' between said first picture and the next closest subsequent P-picture of said bitstream, in a display order of said bitstream;
wherein said bits are allocated for coding said second picture in accordance with said assumed distance M'.

28. The method of claim 27, comprising the further steps of:

determining an actual distance M" between said first picture and the next closest subsequent P-picture, in said display order, according to said picture type of said second picture; and

adjusting said assumed length N' of said GOP in accordance with an actual distance M=M" to provide an adjusted assumed length N"; wherein:

said adjusting step (e) is responsive to said adjusted assumed length N'.

29. The method of claim 28, wherein:

said assumed length N' of said particular GOP is adjusted in said adjusting step thereof by a factor, M(N'/M'), to provide said adjusted assumed length N".

30. The method of claim 28 or 29, wherein said adjusting step (e) comprises the step of:

adjusting said assumed bit budget by a factor(N"-N')*"bit rate"/"frame rate", where "bit rate" is a bit rate of said bitstream, and "frame rate" is a frame rate of said bitstream.

31. The method of one of claims 28 to 30, comprising the further step of:

allocating bits for coding the remaining pictures in said particular GOP following said second picture in accordance with said adjusted bit budget.

32. The method of one of claims 28 to 31, comprising the further step of:

determining if N">N, an actual length of said particular GOP, after coding N" pictures in said particular GOP; and, if N">N:

further adjusting said adjusted bit budget according to N and N"; and

allocating bits for coding pictures in said bitstream following the N" coded pictures according to said further adjusted bit budget.

33. The method of one of claims 28 to 32, comprising the further steps of:

determining if N"<N, an actual length of said particular GOP, after coding N" pictures in said particular GOP; and, if N"<N:

further adjusting said adjusted bit budget according to M=M"; and

allocating bits for coding pictures in said bitstream following the N" coded pictures according to said further adjusted bit budget.

34. The method of one of claims 13 to 33, comprising the further step of:

allocating bits for coding a GOP following said particular GOP in accordance with said adjusted bit budget.

35. The method of one of claims 13 to 34, comprising the further step of:

accounting for changes in an actual distance M" between I-pictures and/or P-pictures in said particular GOP by determining an actual distance M" between an I-picture or P-picture that follows said first picture in said particular GOP, in a display order of said bitstream, and a subsequent I-picture or P-picture in said particular GOP; and

adjusting said assumed bit budget in accordance with said actual distance M".

36. An apparatus for allocating bits for coding pictures in a bitstream received at a digital video transcoder, comprising:

(a) means for providing an assumed distance M' between a first picture of said bitstream and the next closest subsequent P-picture of said bitstream in a display order of said bitstream;

(b) means for providing an assumed bit budget as a function of said assumed distance M';

(c) means for coding said first picture in accordance with said assumed bit budget;

(d) means for determining a picture type of a second picture that immediately follows said first picture in said bitstream in an encoding order of said bit stream;

(e) means for adjusting said assumed bit budget according to said picture type of said second picture; and

(f) means for allocating bits for coding said second picture in accordance with said adjusted bit budget.

37. The apparatus of claim 36, wherein:

said assumed bit budget is proportional to said assumed distance and a bit rate of said bitstream, and inversely proportional to a frame rate of said bitstream.

38. The apparatus of claim 36 or 37, wherein:

said pictures in said bitstream form a progressive refresh sequence.

39. The apparatus of one of claims 36 to 38, wherein said first picture is a P-picture and said second picture is a B-picture, further comprising:

means for determining an actual distance M" between said first picture and said next closest subsequent P-picture, in said display order, according to said picture type of said second picture.

40. The apparatus of claim 39, wherein:

said actual distance M" is determined according to a difference between a temporal reference of said second picture, in said display order, and a temporal reference of said first picture, in said display order, plus one picture.

41. The apparatus of claim 39 or 40, further comprising:

means for allocating bits for coding the remaining M"-1 pictures following said first picture in accordance with said adjusted bit budget.

42. The apparatus of one of claims 39 to 41, wherein:

said adjusting means (e) comprises means for adjusting said assumed bit budget by +(M"-M')*"bit rate"/"frame rate", where "bit rate" is a bit rate of said bitstream, and "frame rate" is a frame rate of said bitstream.

43. The apparatus of one of claims 36 to 42, wherein said first and second pictures are P-pictures, further comprising:

means for determining an actual distance M" between said first picture and said next closest subsequent P-picture, in said display order, according to said picture type of said second picture.

44. The apparatus of claim 43, wherein:

said actual distance M" is determined according to a difference between a temporal reference of said second

picture, in said display order, and a temporal reference of said first picture, in said display order.

45. The apparatus of claim 43 or 44, wherein said adjusting means (e) comprises:

means for adjusting said assumed bit budget by - (M'-1)*"bit rate"/"frame rate", where "bit rate" is a bit rate of said bitstream, and "frame rate" is a frame rate of said bitstream.

46. The apparatus of one of claims 43 to 45, further comprising:

means for allocating bits for coding the remaining M"-1 pictures following said first picture in accordance with said adjusted bit budget.

47. The apparatus of one of claims 36 to 45, further comprising:

means for allocating bits for coding a series of M" pictures following an initial M" pictures that includes said first and second pictures in said bitstream in accordance with said adjusted bit budget.

48. An apparatus for allocating bits for coding a sequence of pictures in a bitstream received at a digital video transcoder, comprising:

(a) means for providing an assumed length N' of a particular group of pictures (GOP) of said bitstream;

(b) means for providing an assumed bit budget as a function of said assumed length N';

(c) means for allocating bits for coding a first picture of said particular GOP in accordance with said assumed bit budget;

(d) means for determining a picture type of a second picture that immediately follows said first picture;

(e) means for adjusting said assumed bit budget according to said picture type of said second picture; and

(f) means for allocating bits for coding said second picture in accordance with said adjusted bit budget.

49. The apparatus of claim 48, wherein:

said assumed bit budget is proportional to said assumed length N' and a bit rate of said bitstream, and inversely proportional to a frame rate of said bitstream.

50. The apparatus of claim 48 or 49, wherein said first and second pictures are I-pictures, further comprising:

means for determining an actual distance M" between said first picture and the next closest subsequent I-picture of said bitstream according to the picture type of said second picture; and

means for adjusting said assumed bit budget in accordance with said actual distance M".

51. The apparatus of claim 50, wherein:

said actual distance M" is determined according to a difference between a temporal reference of said second picture, in a display order of said bitstream, and a temporal reference of said first picture, in said display order.

52. The apparatus of one of claims 48 to 51, wherein said adjusting means (e) comprises:

means for adjusting said assumed bit budget by - (N'-1)*"bit rate"/"frame rate", where "bit rate" is a bit rate of said bitstream, and "frame rate" is a frame rate of said bitstream.

53. The apparatus of one of claims 48 to 52, further comprising:

means for allocating bits for coding the remaining pictures in said particular GOP following said second picture in accordance with said adjusted bit budget.

54. The apparatus of one of claims 48 to 53, further comprising:

means for providing an assumed distance M' between said first picture and the next closest subsequent P-picture of said bitstream, in a display order of said bitstream;
wherein said bits are allocated for coding said first picture by said coding means (c) in accordance with said assumed distance M'.

55. The apparatus of claim 54, wherein said first picture is an I-picture and said second picture is a P-picture, further comprising:

means for determining an actual distance M" between said first picture and the next closest subsequent P-picture, in said display order, according to said picture type of said second picture; and

means for adjusting said assumed length N' of said GOP in accordance with said actual distance M" to provide an adjusted assumed length N"; wherein:

said adjusting means (e) is responsive to said adjusted assumed length N'.

56. The apparatus of claim 55, wherein:

said assumed length N' of said particular GOP is adjusted by said adjusting means thereof by a factor, M(N'/M'), to provide said adjusted assumed length N", where M=M".

57. The apparatus of claim 55 or 56, wherein said adjusting means (e) comprises:

means for adjusting said assumed bit budget by a factor (N"-N')*"bit rate"/"frame rate", where "bit rate" is a bit rate of said bitstream, and "frame rate" is a frame rate of said bitstream.

58. The apparatus of claim 57, wherein said adjusting means (e) comprises:

means for adjusting said assumed bit budget by a factor, (M"-1')*"bit rate"/"frame rate", where "bit rate" is a bit rate of said bitstream, and "frame rate" is a frame rate of said bitstream.

59. The apparatus of one of claims 55 to 58, further comprising:

means for allocating bits for coding the remaining pictures in said particular GOP following said second picture in accordance with said adjusted bit budget.

60. The apparatus of one of claims 55 to 59, further comprising:

means for determining if N">N, an actual length of said particular GOP, after coding N" pictures in said particular GOP;

means for further adjusting said adjusted bit budget according to N and N" if N">N; and

means for allocating bits for coding pictures in said bitstream following the N" coded pictures according to said further adjusted bit budget if N">N.

61. The apparatus of one of claims 55 to 59, further comprising:

means for determining if N"<N, an actual length of said particular GOP, after coding N" pictures in said particular GOP;

means for further adjusting said adjusted bit budget according to M=M" if N"<N; and

means for allocating bits for coding pictures in said bitstream following the N" coded pictures according to said further adjusted bit budget if N"<N.

62. The apparatus of claim 48, wherein said first picture is an I-picture and said second picture is a B-picture, further comprising:

means for providing an assumed distance M' between said first picture and the next closest subsequent P-picture of said bitstream, in a display order of said bitstream;
wherein said bits are allocated for coding said second picture in accordance with said assumed distance M'.

63. The apparatus of claim 62, further comprising:

means for determining an actual distance M" between said first picture and the next closest subsequent P-picture, in said display order, according to said picture type of said second picture; and

means for adjusting said assumed length N' of said GOP in accordance with an actual distance M=M" to provide an adjusted assumed length N"; wherein:

said adjusting means (e) is responsive to said adjusted assumed length N'.

64. The apparatus of claim 63, wherein:

said assumed length N' of said particular GOP is adjusted by said adjusting means thereof by a factor, M(N'/M'), to provide said adjusted assumed length N".

65. The apparatus of claim 63 or 64, wherein said adjusting means (e) comprises:

means for adjusting said assumed bit budget by a factor (N"-N')*"bit rate"/"frame rate", where "bit rate" is a bit rate of said bitstream, and "frame rate" is a frame rate of said bitstream.

66. The apparatus of one of claims 63 to 65, further comprising:

means for allocating bits for coding the remaining pictures in said particular GOP following said second picture in accordance with said adjusted bit budget.

67. The apparatus of one of claims 63 to 66, further comprising:

means for determining if N">N, an actual length of said particular GOP, after coding N" pictures in said particular GOP;

means for further adjusting said adjusted bit budget according to N and N" if N">N; and

means for allocating bits for coding pictures in said bitstream following the N" coded pictures according to said further adjusted bit budget if N">N.

68. The apparatus of one of claims 63 to 66, further comprising:

means for determining if N"<N, an actual length of said particular GOP, after coding N" pictures in said particular GOP;

means for further adjusting said adjusted bit budget according to M=M" if N"<N; and

means for allocating bits for coding pictures in said bitstream following the N" coded pictures according to said further adjusted bit budget if N"<N.

69. The apparatus of one of claims 48 to 68, further comprising:

means for allocating bits for coding a GOP following said particular GOP in accordance with said adjusted

bit budget.

70. The apparatus of one of claims 48 to 69, further comprising:

means for accounting for changes in an actual distance M" between I-pictures and/or P-pictures in said particular GOP by determining an actual distance M" between an I-picture or P-picture that follows said first picture in said particular GOP, in a display order of said bitstream, and a subsequent I-picture or P-picture in said particular GOP; and

means for adjusting said assumed bit budget in accordance with said actual distance M".

71. A method for allocating bits for coding a sequence of pictures in a bitstream received at a digital video transcoder, comprising the steps of:

(a) providing a length of groups of pictures (GOPs) of said bitstream;

(b) providing a distance between each I-picture and a next successive P-picture in said GOPs;

(c) maintaining a count of a number of remaining I-, P- and B-pictures, respectively, in each GOP as each picture in said GOPs is coded;

(d) providing an assumed bit budget in response to said steps (a), (b) and (c) for coding pictures in said GOPs;

(e) verifying, and adjusting, if necessary, said length after each I-picture in said GOPs is coded;

(f) verifying, and adjusting, if necessary, said distance after each I-picture and/or P-picture in said GOPs is coded;

(g) verifying, and adjusting, if necessary, the count of the number of remaining I-, P- and B-pictures, in response to said steps (e) and (f); and

(g) adjusting, if necessary, said assumed bit budget in response to said steps (e), (f) and (g).

72. The method of claim 71, wherein:

said distance is assumed when initially coding said GOPs and is corrected to an actual distance within a time for coding two pictures in said bitstream.

73. The method of claim 71 or 72, wherein:

said length is assumed when initially coding said GOPs and is corrected to an actual length within a time for coding one GOP in said bitstream.

74. An apparatus for allocating bits for coding a sequence of pictures in a bitstream received at a digital video transcoder, comprising:

(a) means for providing a length of groups of pictures (GOPs) of said bitstream;

(b) means for providing a distance between each I-picture and a next successive P-picture in said GOPs;

(c) means for maintaining a count of a number of remaining I-, P- and B-pictures, respectively, in each GOP as each picture in said GOPs is coded;

(d) means for providing an assumed bit budget in response to said means (a), (b) and (c) for coding pictures in said GOPs;

(e) means for verifying, and adjusting, if necessary, said length after each I-picture in said GOPs is coded;

(f) means for verifying, and adjusting, if necessary, said distance after each I-picture and/or P-picture in said GOPs is coded;

(g) means for verifying, and adjusting, if necessary, the count of the number of remaining I-, P- and B-pictures, in response to said means (e) and (f); and

(g) means for adjusting, if necessary, said assumed bit budget in response to said means (e), (f) and (g).

75. The apparatus of claim 74, wherein:

said distance is assumed when initially coding said GOPs and is corrected to an actual distance within a time for coding two pictures in said bitstream.

76. The apparatus of claim 74 or 75, wherein:

said length is assumed when initially coding said GOPs and is corrected to an actual length within a time for coding one GOP in said bitstream.

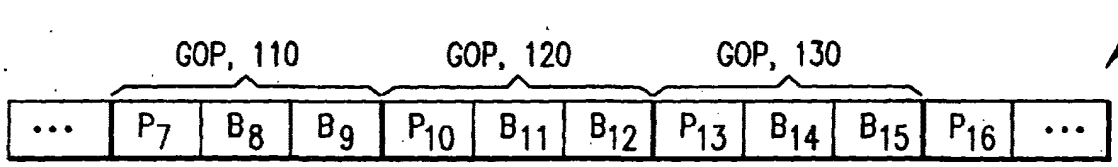


FIG.1a

GOP(M) DISPLAY ORDER–CONTINUING GOP

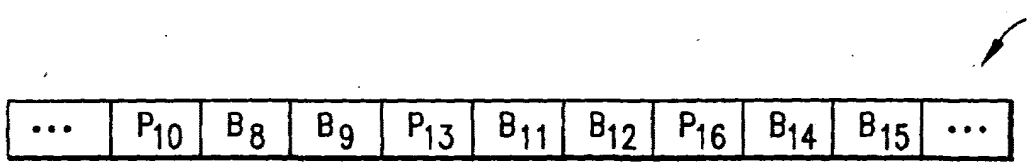


FIG.1b

GOP(M) ENCODING ORDER–CONTINUING GOP

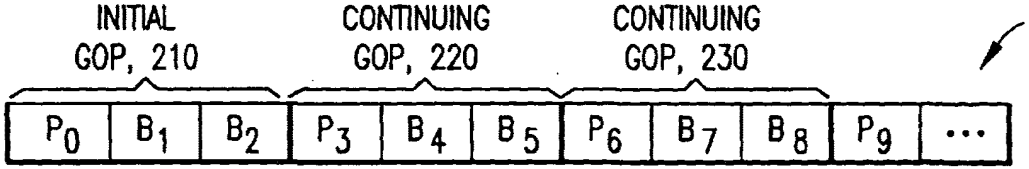


FIG.2a

GOP(M) DISPLAY ORDER–INITIAL GOP

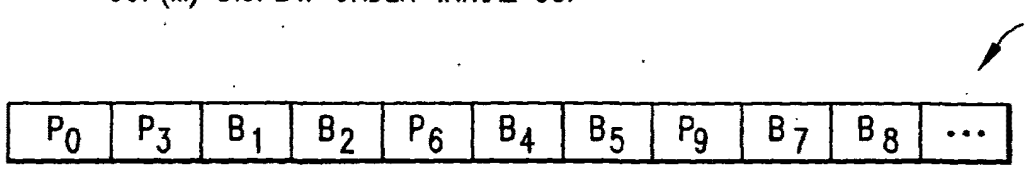


FIG.2b

GOP(M) ENCODING ORDER–INITIAL GOP

EP 1 005 232 A2

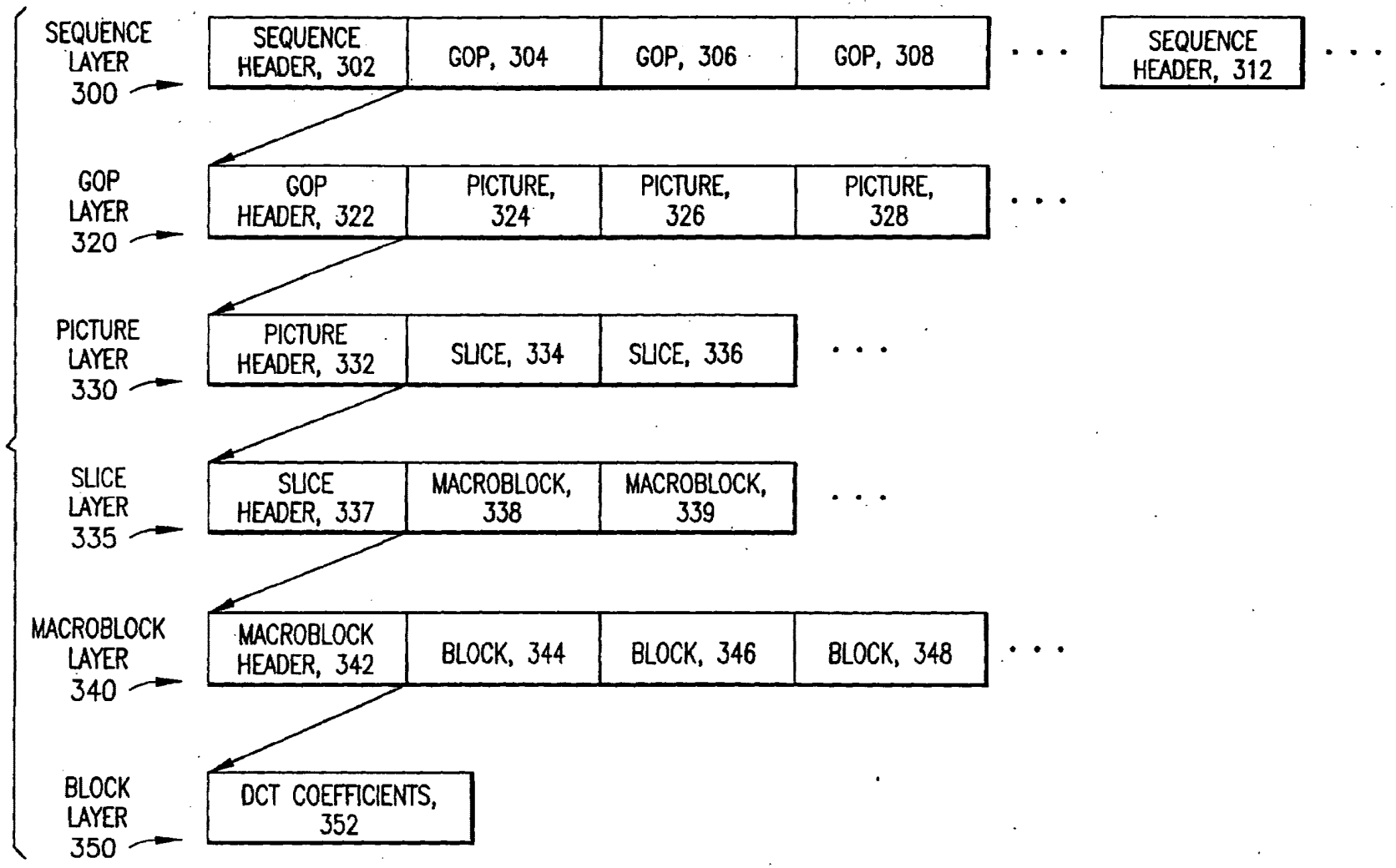

SEQUENCE LAYER 300
SEQUENCE HEADER, 302
GOP, 304
GOP, 306
GOP, 308
SEQUENCE HEADER, 312
GOP LAYER 320
GOP HEADER, 322
PICTURE, 324
PICTURE, 326
PICTURE, 328
PICTURE LAYER 330
PICTURE HEADER, 332
SLICE, 334
SLICE, 336
SLICE LAYER 335
SLICE HEADER, 337
MACROBLOCK, 338
MACROBLOCK, 339
MACROBLOCK LAYER 340
MACROBLOCK HEADER, 342
BLOCK, 344
BLOCK, 346
BLOCK, 348
BLOCK LAYER 350
DCT COEFFICIENTS, 352

FIG.3

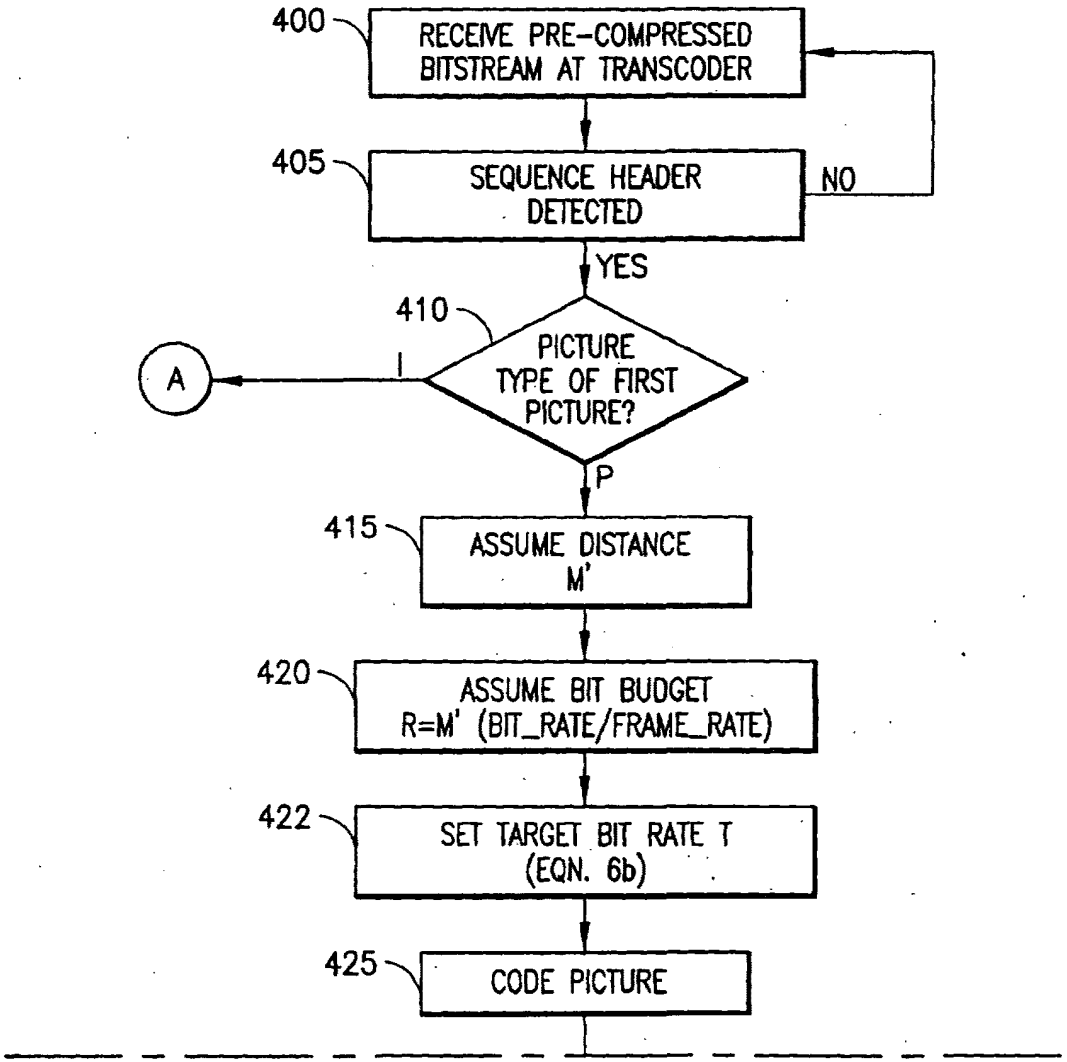


FIG.4a-1
FIG.4a-2

FIG.4a

FIG.4a-1

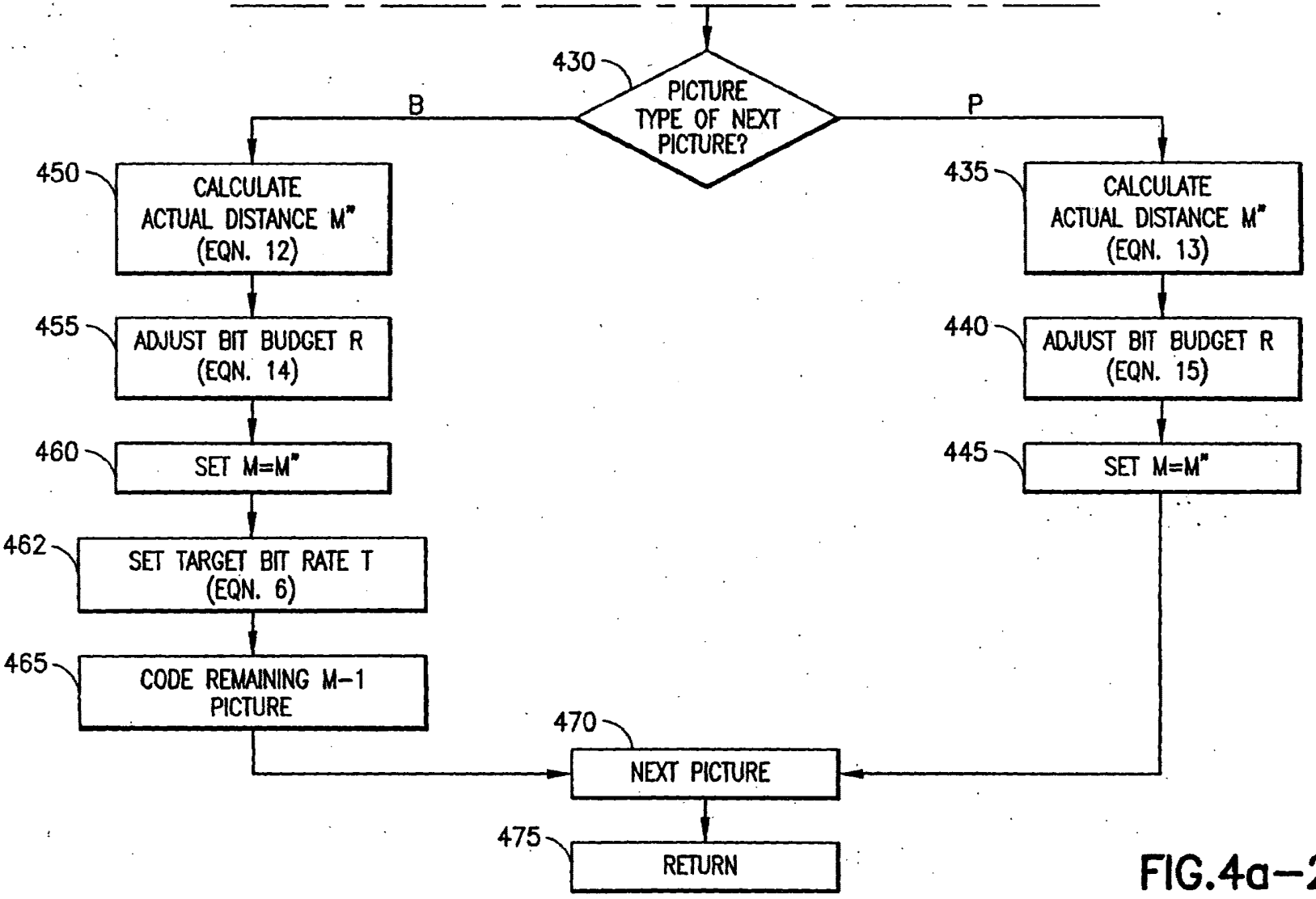

430
PICTURE TYPE OF NEXT PICTURE?
B
P
450
CALCULATE ACTUAL DISTANCE M" (EQN. 12)
455
ADJUST BIT BUDGET R (EQN. 14)
460
SET M=M"
462
SET TARGET BIT RATE T (EQN. 6)
465
CODE REMAINING M-1 PICTURE
435
CALCULATE ACTUAL DISTANCE M" (EQN. 13)
440
ADJUST BIT BUDGET R (EQN. 15)
445
SET M=M"
470
NEXT PICTURE
475
RETURN

FIG.4a-2

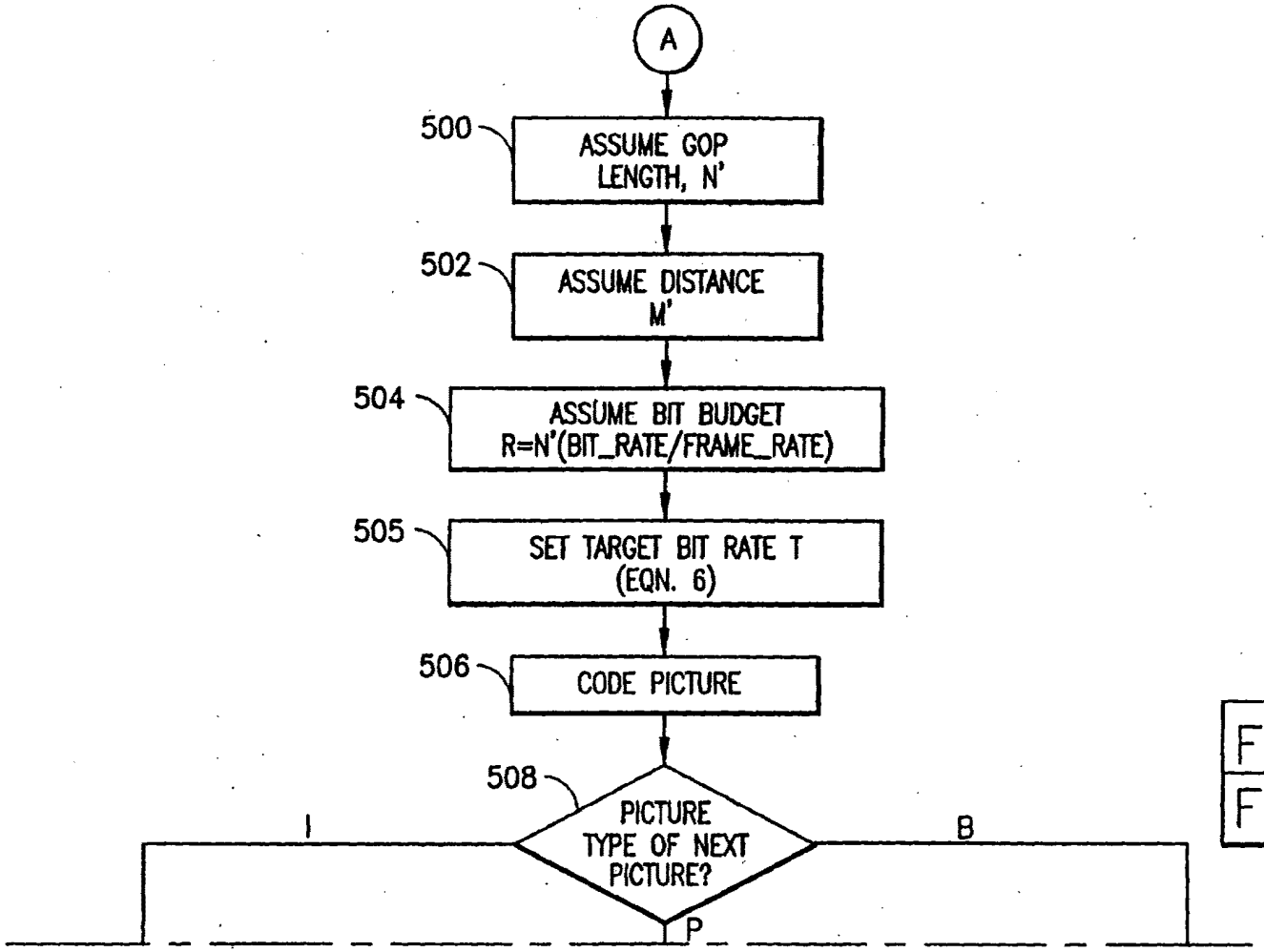


FIG.4b-1

FIG.4b-1
FIG.4b-2

FIG.4b

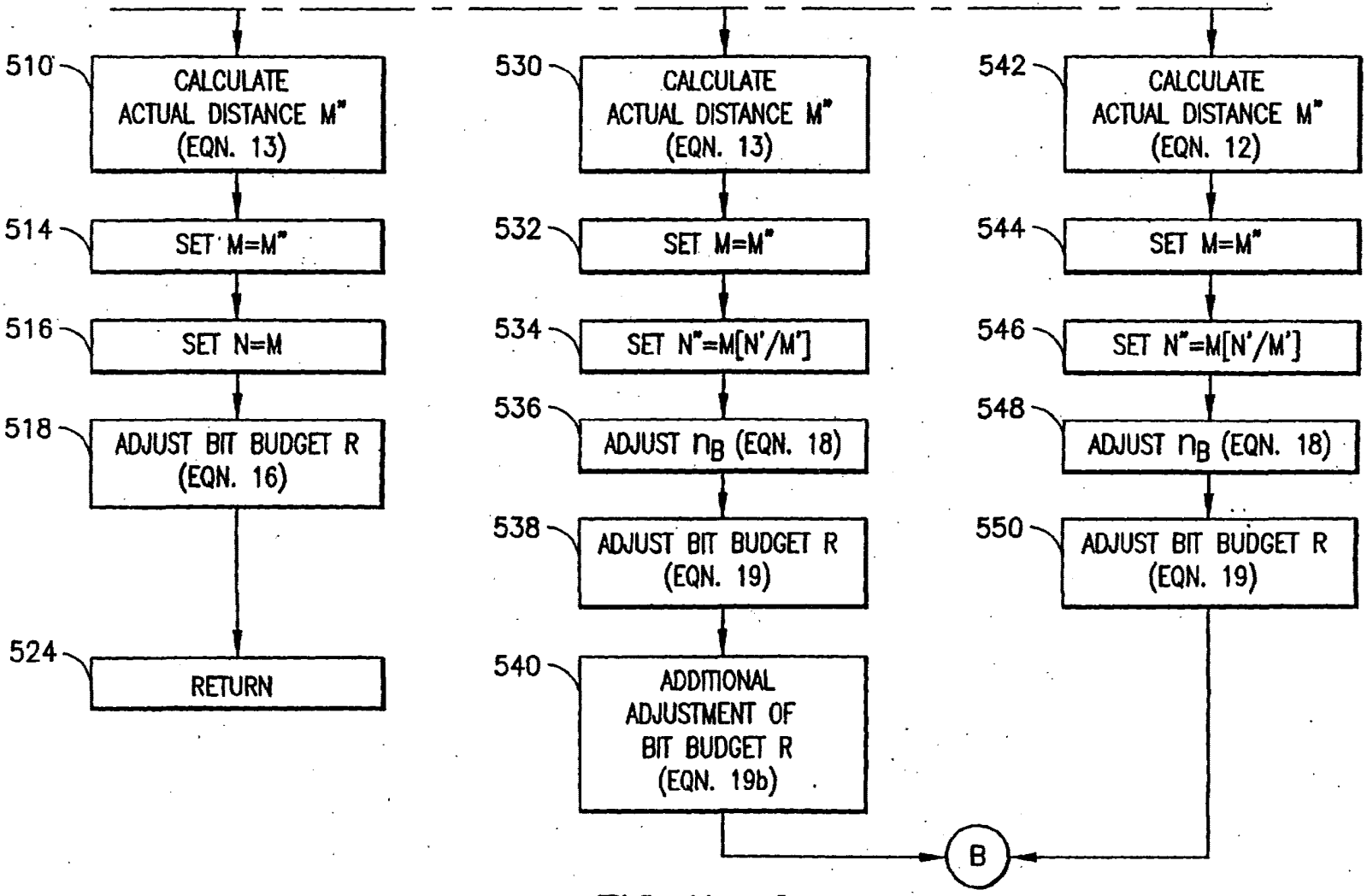


FIG.4b–2

EP 1 005 232 A2

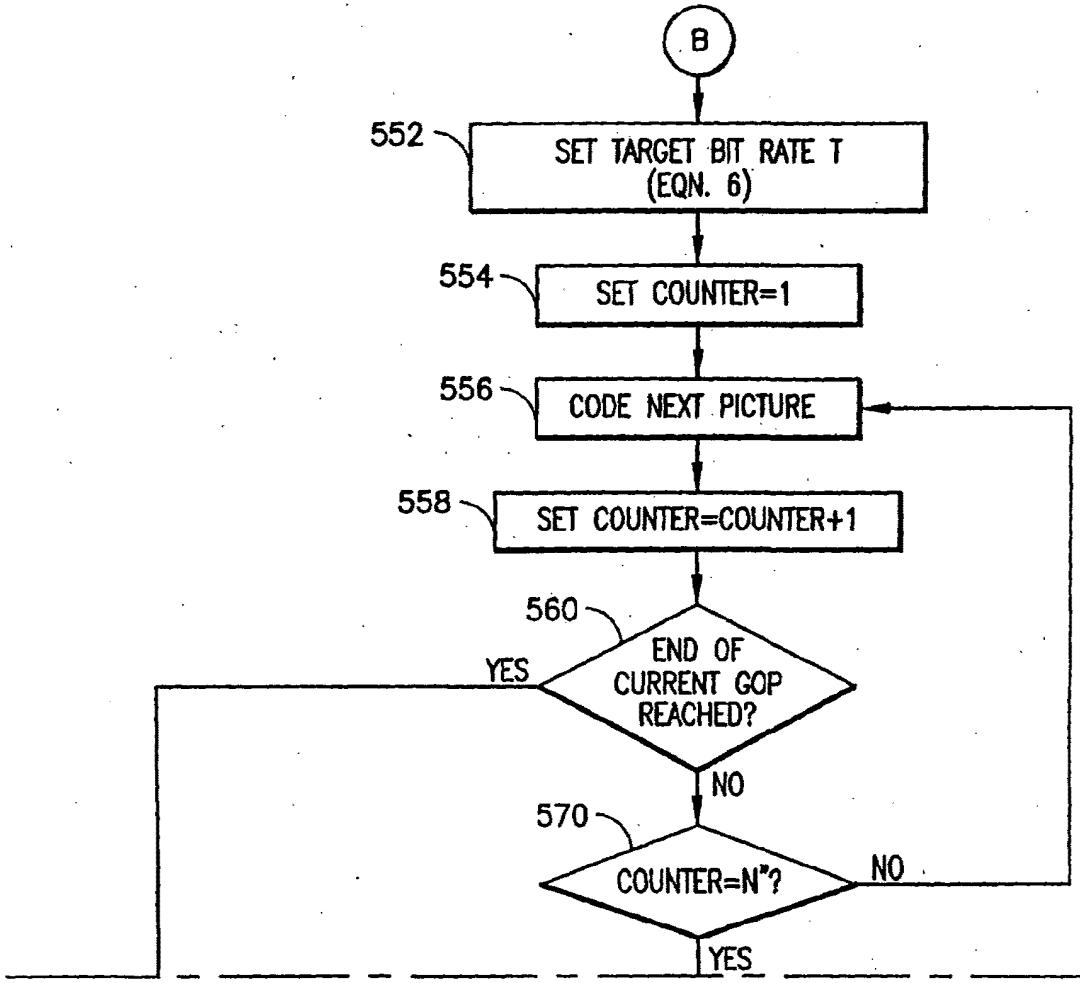


FIG.4c-1

FIG.4c-1
FIG.4c-2

FIG.4c

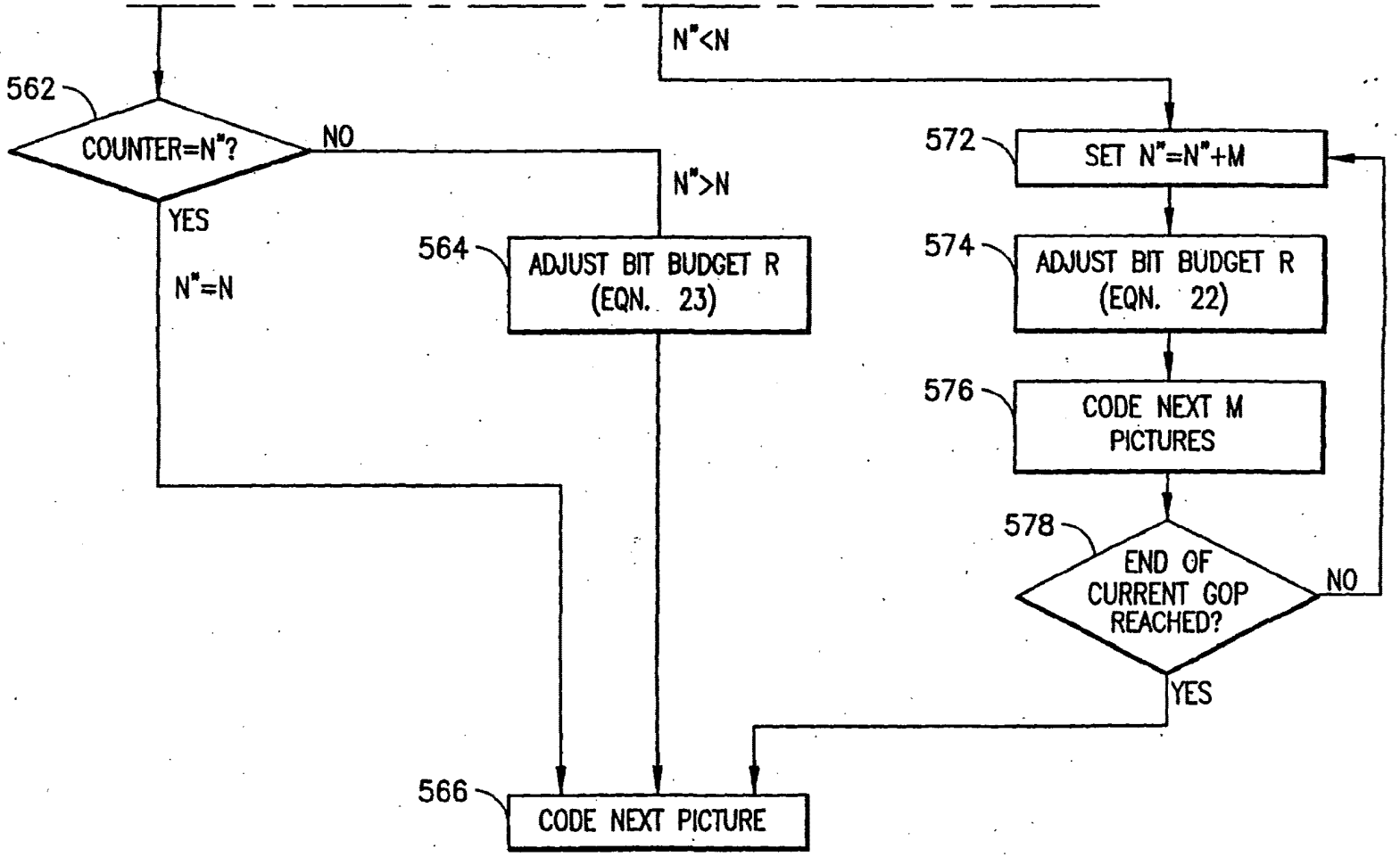

N"<N
562
COUNTER=N"?
NO
N">N
YES
N"=N
564
ADJUST BIT BUDGET R
(EQN. 23)
572
SET N"=N"+M
574
ADJUST BIT BUDGET R
(EQN. 22)
576
CODE NEXT M
PICTURES
578
END OF
CURRENT GOP
REACHED?
NO
YES
566
CODE NEXT PICTURE

FIG.4c–2

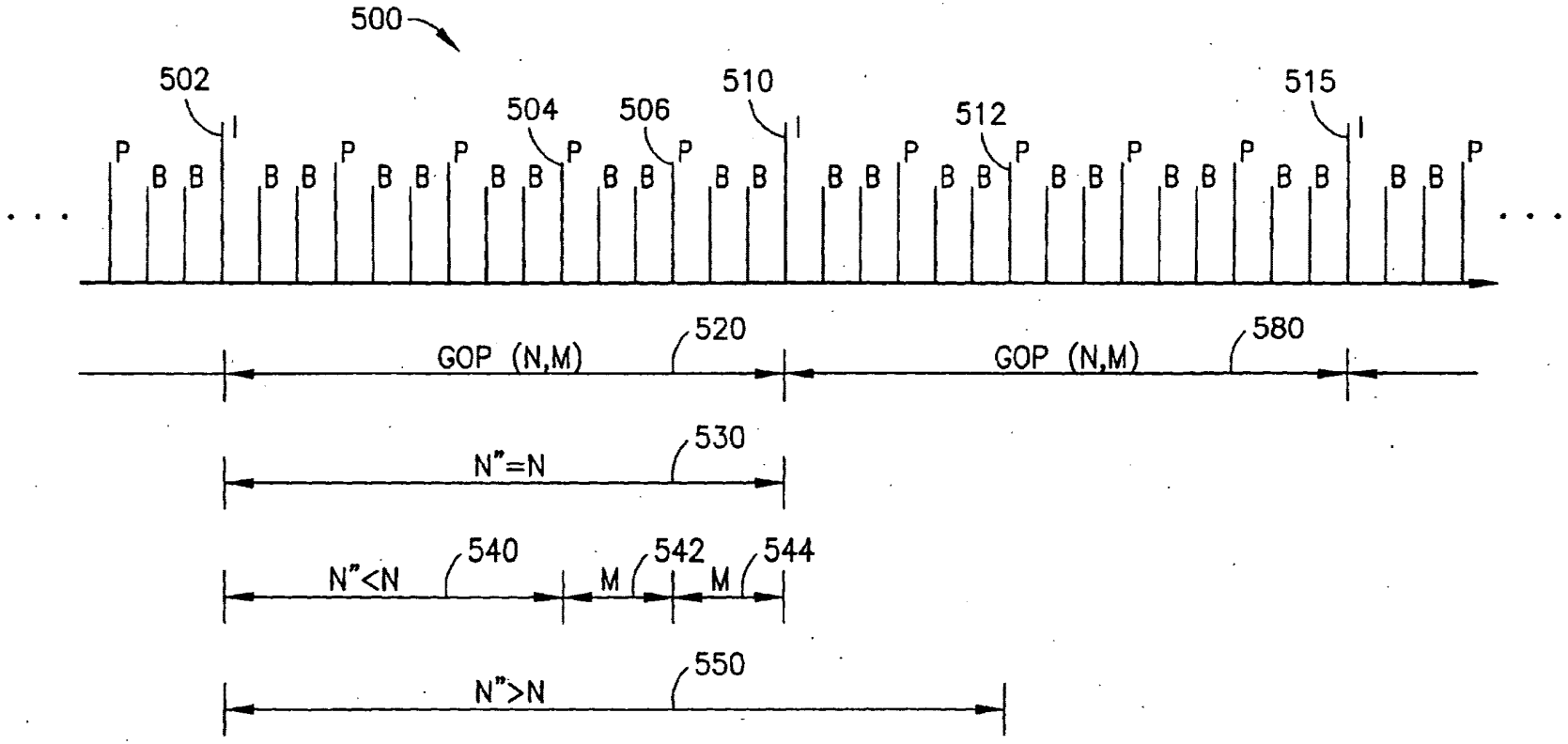


FIG.5

EP 1 005 232 A2

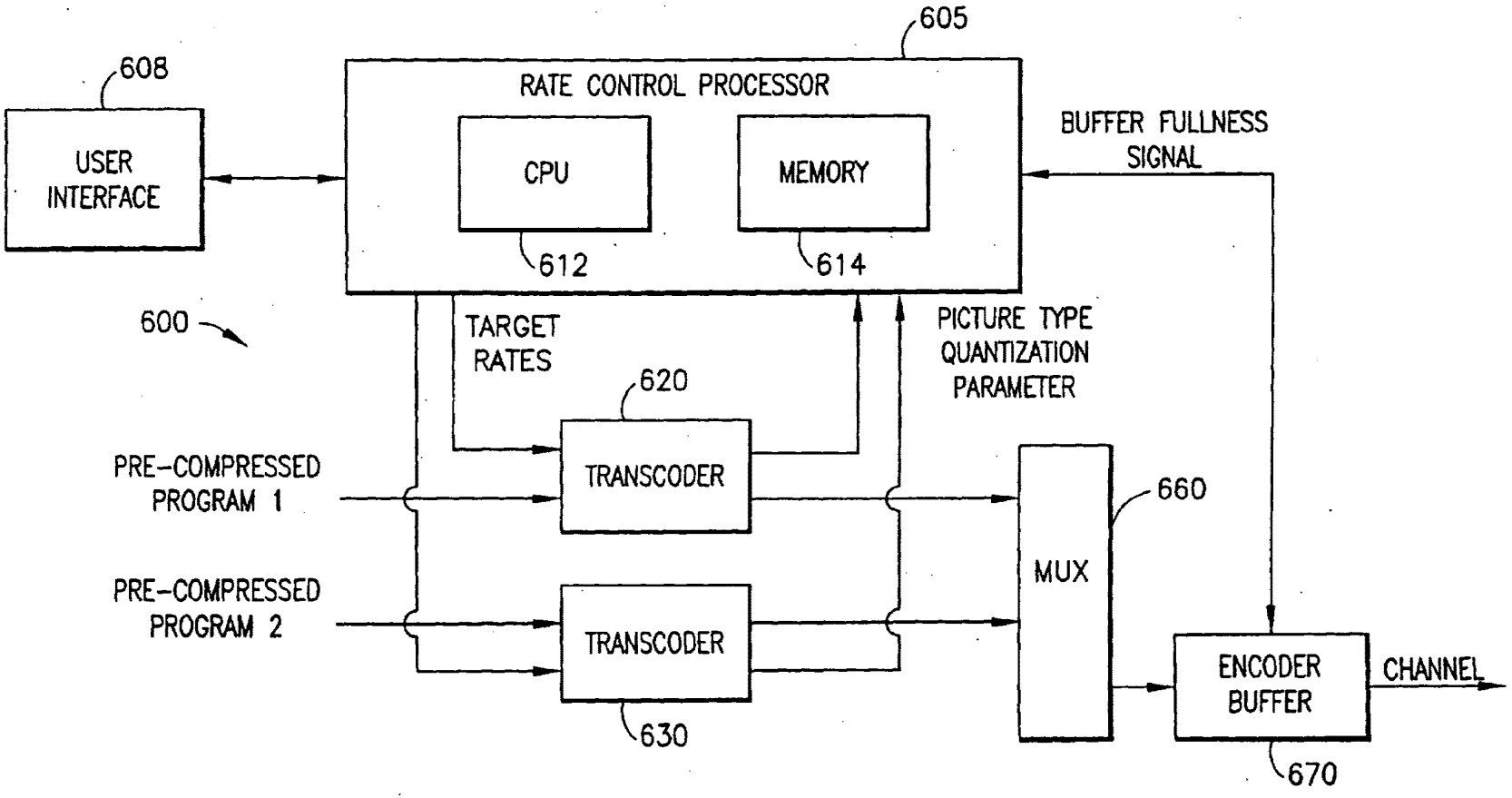

605
608
RATE CONTROL PROCESSOR
USER INTERFACE
CPU
MEMORY
BUFFER FULLNESS SIGNAL
612
614
600
TARGET RATES
PICTURE TYPE QUANTIZATION PARAMETER
620
PRE-COMPRESSED PROGRAM 1
TRANSCODER
660
MUX
PRE-COMPRESSED PROGRAM 2
TRANSCODER
ENCODER BUFFER
CHANNEL
630
670

FIG.6

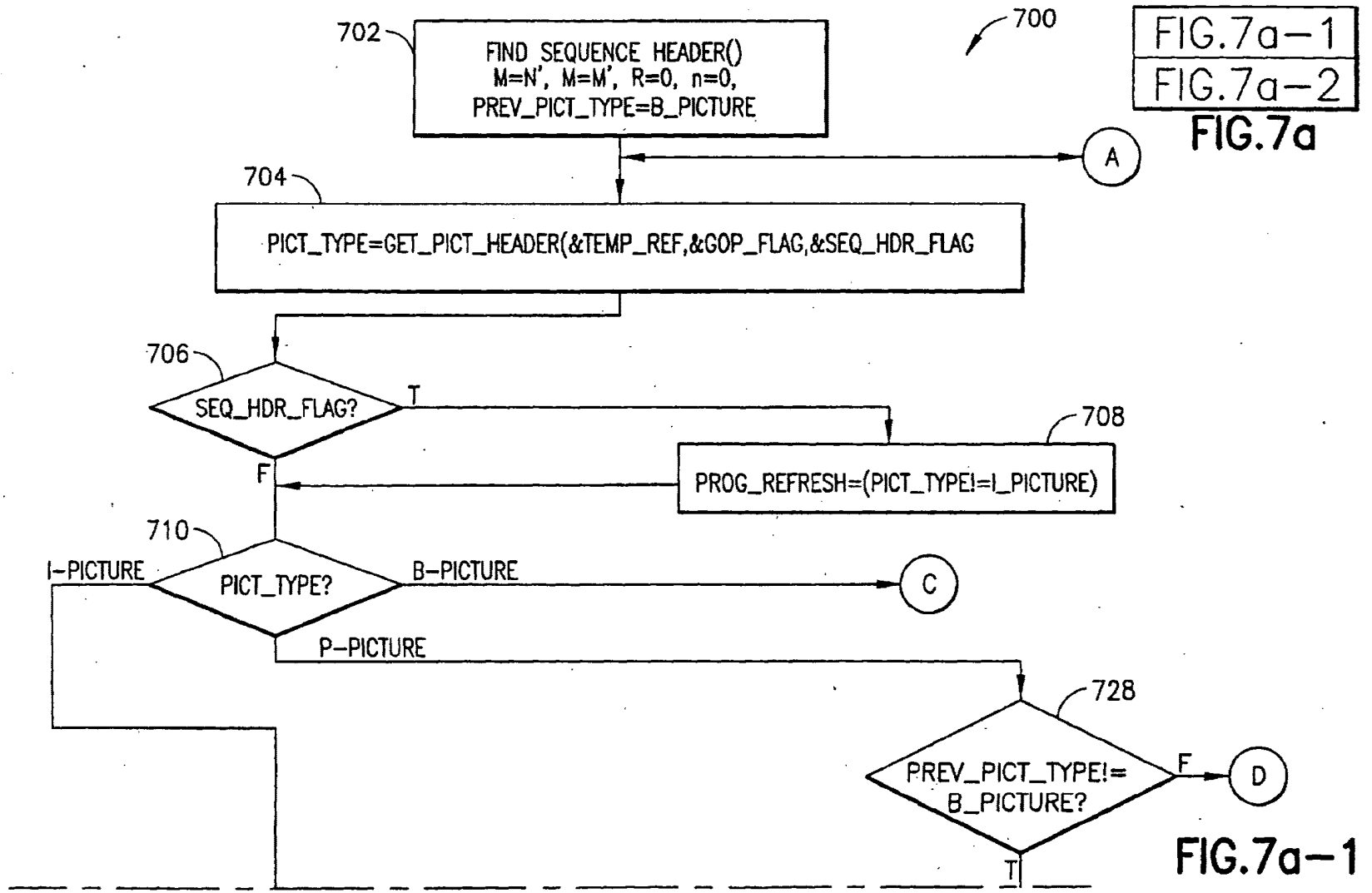


FIG.7a-1
FIG.7a-2
FIG.7a

FIG.7a-1

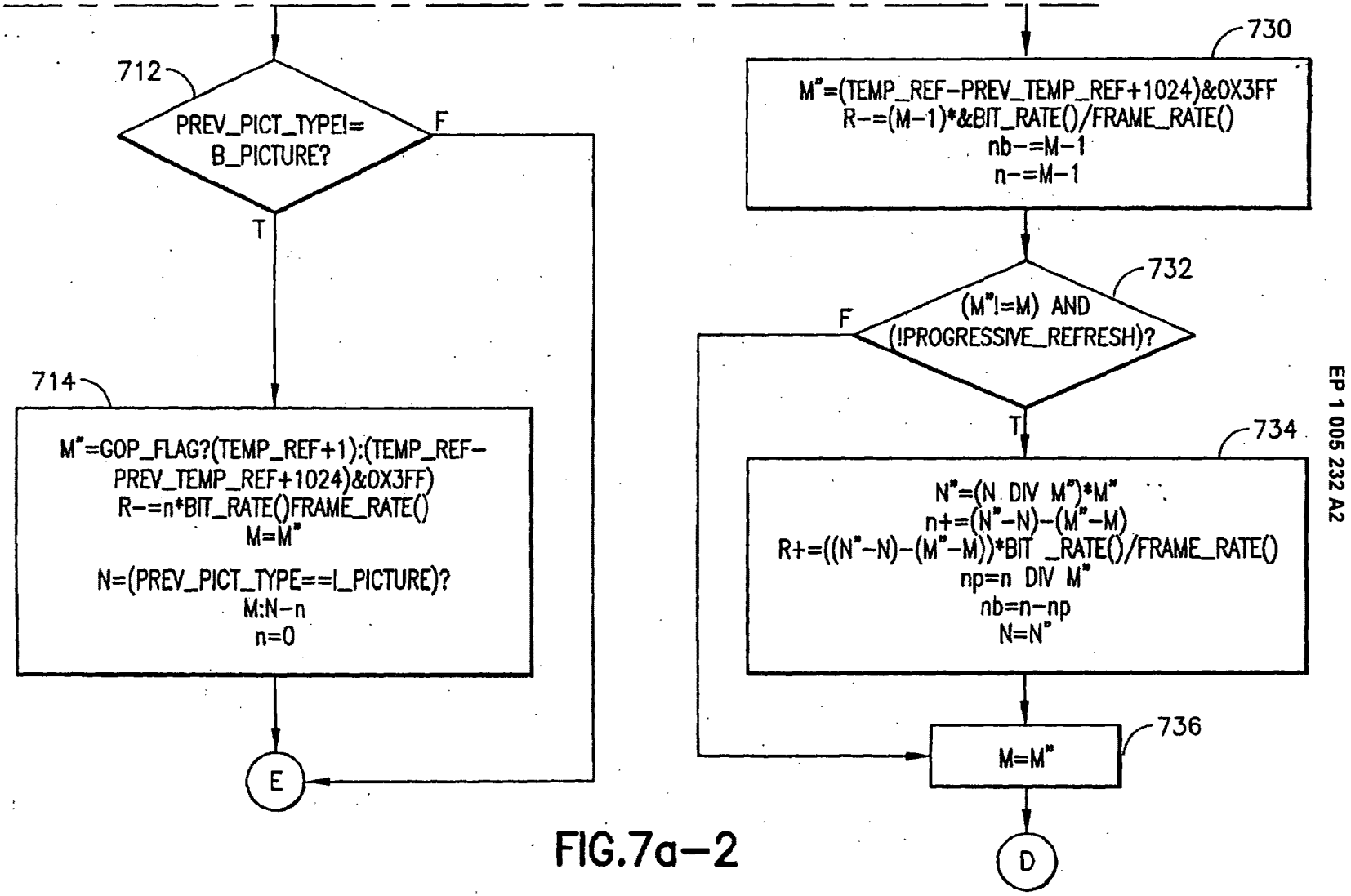


FIG.7a-2

EP 1 005 232 A2

OIPE
AUG 18 2006
PATENT & TRADEMARK OFFICE

IFW

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant:	HANNUKSELA	Examiner:	Unassigned
Serial No.:	11/338,934	Group Art Unit:	2613
Filed:	January 25, 2006	Docket No.:	TPO.004.C2
Title:	GROUPING OF IMAGE FRAMES IN VIDEO CODING		

OIPE
AUG 18 2006
PATENT & TRADEMARK OFFICE

CERTIFICATE UNDER 37 CFR 1.8: The undersigned hereby certifies that this correspondence and the papers, as described hereinabove, are being deposited in the United States Postal Service, as first class mail, in an envelope addressed to: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450 on August 15, 2006.

By: Tracey M. Dotter
Tracey M. Dotter

SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT (37 C.F.R. §1.97(b))

Mail Stop Amendment
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

With regard to the above-identified application, the item of information listed on the enclosed Form 1449 is brought to the attention of the Examiner.

This statement should be considered because it is submitted either within three months of the filing date or before the first Office Action on the merits of the above-identified application. Accordingly, no fee is due for consideration of the items listed on the enclosed Form 1449.

In accordance with 37 C.F.R. §1.98(a)(2), and the 5 August 2003 Official Gazette Notice, only a copy of each foreign document or other non-patent published document listed on the enclosed Form 1449 is provided.

Please note that any notations or markings on the attached documents do not reflect particular relevance, or lack thereof, to the present application, nor were they necessarily made by anyone affiliated with the prosecution of the present application.

No representation is made that a reference is "prior art" within the meaning of 35 U.S.C. §§102 and 103 and Applicant reserves the right, pursuant to 37 C.F.R. §1.131 or otherwise, to establish that the reference(s) are not "prior art." Moreover, Applicant does not represent that a reference has been thoroughly reviewed or that any relevance of any portion of a reference is intended.

Consideration of the items listed is respectfully requested. Pursuant to the provisions of MPEP §609, it is requested that the Examiner return a copy of the attached Form 1449, marked as being considered and initialed by the Examiner, to the undersigned with the next official communication.

Respectfully submitted,

HOLLINGSWORTH & FUNK, LLC
8009 34th Avenue South, Suite 125
Minneapolis, MN 55425
952.854.2700

Date: August 15, 2006

By: E. M. Nichols
Erin M. Nichols
Reg. No. 57,125

FORM 1449* INFORMATION DISCLOSURE STATEMENT IN AN APPLICATION (Use several sheets if necessary)	Docket Number: TPO.004.C2	Application Number: 11/338,934
	Applicant: HANNUKSELA	
	Filing Date: 1/25/2006	Group Art Unit: 2613

OIPE AUG 18 2006 PATENT & TRADEMARK OFFICE IAP19

U.S. PATENT DOCUMENTS

EXAMINER INITIAL	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE

FOREIGN PATENT DOCUMENTS

	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	TRANSLATION	
						YES	NO

OTHER DOCUMENTS (Including Author, Title, Date, Pertinent Pages, Etc.)

	1998	Wenger, "Temporal Scalability Using P-pictures for Low-latency Applications", IEEE Workshop, December 1998, pp. 559-564.

EXAMINER	DATE CONSIDERED
EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609; draw line through citation if not in conformance and not considered. Include copy of this form for next communication to the Applicant.	

*Substitute Disclosure Statement Form (PTO-1449) Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

OIPE
NOV 24 2006
IAP43
PATENT & TRADEMARK OFFICE

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant:	HANNUKSELA	Examiner:	Unassigned
Serial No.:	11/338,934	Group Art Unit:	2613
Filed:	January 25, 2006	Docket No.:	TPO.004.C2
Title:	GROUPING OF IMAGE FRAMES IN VIDEO CODING		

CERTIFICATE UNDER 37 CFR 1.8: The undersigned hereby certifies that this correspondence and the papers, as described hereinabove, are being deposited in the United States Postal Service, as first class mail, in an envelope addressed to: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450 on November 21, 2006.

By: Rennae Johnson

SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT (37 C.F.R. §1.97(b))

Mail Stop Amendment
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

With regard to the above-identified application, the item of information listed on the enclosed Form 1449 is brought to the attention of the Examiner.

This statement should be considered because it is submitted either within three months of the filing date or before the first Office Action on the merits of the above-identified application. Accordingly, no fee is due for consideration of the item listed on the enclosed Form 1449.

In accordance with 37 C.F.R. §1.98(a)(2), and the 5 August 2003 Official Gazette Notice, only a copy of each foreign document or other non-patent published document listed on the enclosed Form 1449 is provided.

Please note that any notations or markings on the attached document do not reflect particular relevance, or lack thereof, to the present application, nor were they necessarily made by anyone affiliated with the prosecution of the present application.

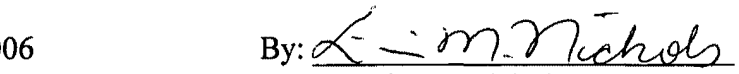
No representation is made that a reference is "prior art" within the meaning of 35 U.S.C. §§102 and 103 and Applicant reserves the right, pursuant to 37 C.F.R. §1.131 or otherwise, to establish that the reference(s) are not "prior art." Moreover, Applicant does not represent that a reference has been thoroughly reviewed or that any relevance of any portion of a reference is intended.

Consideration of the item listed is respectfully requested. Pursuant to the provisions of MPEP §609, it is requested that the Examiner return a copy of the attached Form 1449, marked as being considered and initialed by the Examiner, to the undersigned with the next official communication.

Respectfully submitted,

HOLLINGSWORTH & FUNK, LLC
8009 34th Avenue South, Suite 125
Minneapolis, MN 55425
952.854.2700

Date: November 20, 2006

By: 
Erin M. Nichols
Reg. No. 57,125

O I P E
IAP43
NOV 24 2006
PATENT & TRADEMARK OFFICE

FORM 1449* INFORMATION DISCLOSURE STATEMENT IN AN APPLICATION (Use several sheets if necessary)	Docket Number: TPO.004.C2	Application Number: 11/338,934
	Applicant: HANNUKSELA	
	Filing Date: 1/25/2006	Group Art Unit: 2613

U.S. PATENT DOCUMENTS

EXAMINER INITIAL	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE

FOREIGN PATENT DOCUMENTS

	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	TRANSLATION	
						YES	NO
	RU 2 123 769	20.12.1998	Russia			Abstract Only	

OTHER DOCUMENTS (Including Author, Title, Date, Pertinent Pages, Etc.)

EXAMINER	DATE CONSIDERED
EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609; draw line through citation if not in conformance and not considered. Include copy of this form for next communication to the Applicant.	

*Substitute Disclosure Statement Form (PTO-1449) Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

(19) RU (11) 2 123 769 (13) C1

(51) МПК6 H 04 N 9/79

РОССИЙСКОЕ АГЕНТСТВО
ПО ПАТЕНТАМ И ТОВАРНЫМ ЗНАКАМ

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ РОССИЙСКОЙ ФЕДЕРАЦИИ

(21), (22) Заявка: 94046325/09, 11.04.1994

(30) Приоритет: 09.04.1993 JP P5-105943

(46) Дата публикации: 20.12.1998

(56) Ссылки: US 4723175 A, 02.02.88. DE 3809207 A, 29.09.88. JP 3-217183 A, 24.09.91. JP 3-252284 A, 11.11.91. SU 1434563 A, 30.10.88.

(86) Заявка PCT:
JP 94/00610 (11.04.94)

(71) Заявитель:
Сони Корпорейшн (JP)

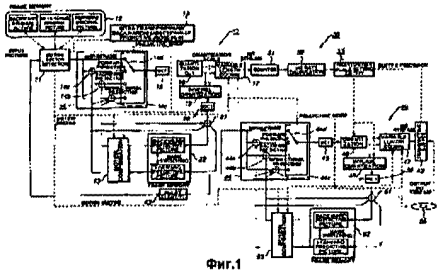
(72) Изобретатель: Цуеси Ода (JP)

(73) Патентообладатель:
Сони Корпорейшн (JP)

(54) СПОСОБЫ И УСТРОЙСТВА КОДИРОВАНИЯ ИЗОБРАЖЕНИЙ И НОСИТЕЛИ ИНФОРМАЦИИ ДЛЯ ЗАПИСИ ИЗОБРАЖЕНИЙ

(57) Реферат:

Электронная схема и последовательность операций, позволяющие кодировать входной видеосигнал путем выполнения, например, прогнозирующего кодирования, ДКП обработки (Дискретное Косинусное Преобразование), квантования данных с фиксированным шагом квантования, неравномерного кодирования (с переменной длиной) для вырабатывания первых кодированных данных и последующего определения (начисления) количества кодов, отводимых (задаваемых) для каждого кадра или каждой группы изображений ГИ на основе учета количества данных в каждый заранее задаваемый отрезок времени, например за время длительности каждого кадра или каждой COP в объеме первых кодированных данных и общего количества данных, пригодных для кодирования входного видеосигнала в каждый заранее задаваемый отрезок времени на основе назначаемого количества кодов, для выработки вторых кодированных данных. Таким образом осуществляется кодирование с переменной скоростью, когда скорость кодирования изменяется в каждый заранее задаваемый отрезок времени. В результате, если даже имеет место последовательность изображений (кадров) сложного рисунка, нет возможности обеспечить столь же большое значение шага квантования для этих изображений, как в обычных устройствах. Благодаря этому можно достигнуть однородного изображения высокого качества в целом в самых различных случаях, что является достигаемым техническим результатом. Кроме того, поскольку вторые кодированные данные имеют переменную скорость, то в случае, когда такие данные регистрируются на носителе записанного изображения, ограниченный объем памяти может быть использован более эффективно, а время записи изображения может быть продлено. Наконец, данные изображения, имеющего высокое качество и однородность во всех возможных случаях, могут быть воспроизведены с носителя записанного изображения. 9 с. и 18 з.п.ф-лы, 12 ил.



Фиг.1

RU 2123769 C1

RU 2123769 C1

(19) **RU** (11) **2 123 769** (13) **C1**

(51) Int. Cl.[6] **H 04 N 9/79**

RUSSIAN AGENCY
FOR PATENTS AND TRADEMARKS

(12) ABSTRACT OF INVENTION

(21), (22) Application: 94046325/09, 11.04.1994

(30) Priority: 09.04.1993 JP P5-105943

(46) Date of publication: 20.12.1998

(86) PCT application:
JP 94/00610 (11.04.94)

(71) Applicant:
Soni Korporejshn (JP)

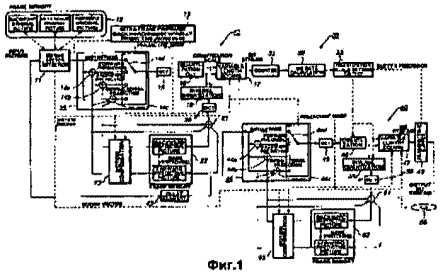
(72) Inventor: **Tsuesi Oda (JP)**

(73) Proprietor:
Soni Korporejshn (JP)

(54) METHOD AND DEVICE FOR ENCODING IMAGES AND INFORMATION MEDIUM FOR STORING IMAGES

(57) Abstract:

FIELD: computer engineering, in particular, for encoding of input signal by running prediction encoding, discrete cosine transform, data sampling with constant quantization step, non-uniform encoding (with alternating length). SUBSTANCE: method involves encoding of first data group and subsequent detection (assigning) code length for allocation (setting) for each frame or each image group using data length in each given time interval, for example, for duration of each frame in first encoded data group and total length of data available for encoding of input video signal in each given time interval using assigned number of codes for generation of second group of encoded data. This results in possibility of alternating length encoding, when encoding rate alters in any given time interval. So, even if sequence of images (frames) for complex picture is given, there is not way to ensure large quantization step for these images as in state-of-the-art devices. This provides possibility to generate uniform image of high quality in various cases. In addition since second group of encoded data has alternating rate, then if such data are recorded onto information medium, a restricted size of buffer memory may be used with greater efficiency and information storage duration may be extended. Finally, image data of high quality and uniformity in all possible cases can be read from medium which stores this image. EFFECT: increased image quality, increased efficiency of buffer memory usage. 28 cl, 12 dwg



Фиг.1

RU 2123769 C1

RU 2123769 C1

Область применения.

Настоящее изобретение относится к способу кодирования подвижного изображения, к устройству кодирования подвижного изображения и к носителю информации в виде записанного подвижного изображения, а более конкретно оно относится к способу и устройству кодирования подвижного изображения, к носителю записанного подвижного изображения, который используется в системе для кодирования и последующего хранения, к видеосигналу подвижного изображения, к регистрации кодированного сигнала на носителе информации, таком как, например, оптический диск, магнитный диск или магнитная лента и т.п., а также к системам для передачи видеосигнала подвижного изображения через передающую среду.

Предпосылки создания изобретения.

До последнего времени в системах передачи видеосигнала подвижного изображения в удаленные точки, как, например, в телевизионных системах передачи с конференцией, в телевизионных телефонных системах и т.п., или в системах регистрации видеосигнала на таких носителях, как оптический диск, магнитный диск или магнитная лента и т.п., или при воспроизведении записанного видеосигнала подвижного изображения, была широко распространена так называемая схема эффективного кодирования при более эффективном использовании трассы передачи видеосигнала (или носителя записи изображения), когда за счет использования корреляции между строками или корреляции между кадрами видеосигнала снижается избыточность по пространственной оси и по временной оси, и передается только наиболее существенная информация, что увеличивает эффективность передачи.

Например, при обработке в виде кодирования в направлении пространственной оси (далее в тексте - внутрикадровое кодирование) корреляция между строками видеосигнала используется так, как показано на фиг. 7. При попытке передать соответствующие изображения PC1, PC2, PC3,... представляющие собой в моменты времени t1, t2, t3,... подвижные изображения, данные изображения, подлежащие обработке и передаче, подвергаются одномерному кодированию, например, в рамках одной и той же строки, или же изображение разбивается, например, на несколько блоков, что позволяет подвергнуть данные изображения в соответствующих блоках двумерному кодированию с тем, чтобы осуществить сжатие данных, что увеличивает эффективность передачи.

Кроме того, при обработке кодированием в направлении оси времени (далее также - внутрикадровое кодирование) внутрикадровая корреляция видеосигнала используется для определения (с помощью так называемого прогнозирующего кодирования) данных изображения PC12, PC23,... имеющих недостаток элементов (так называемых прогнозированных ошибок) в каждом из пикселей данных изображений, общих для смежных изображений PC1, PC2 и PC3, в последовательности этих данных PC12, PC23,... подлежащих передаче, осуществляя, таким образом, сжатие данных, а следовательно, и увеличивая эффективность передачи.

Таким образом, по сравнению со случаем, когда передаются все данные изображений PC1, PC2, PC3,... в данном случае видеосигнал может быть передан на основе значительно меньшего количества данных.

Кроме того, в процессе прогнозирующего кодирования во время описанного выше внутрикадрового кодирования применяется также прогнозируемая компенсация движения, например, в объеме массивов, состоящих из макроблоков, с тем, чтобы еще более увеличить эффективность. Например, когда персонаж в центральной части изображения движется, такое движение объекта на изображении обнаруживается с целью коррекции положения данных изображения, используемых для прогнозирования в предыдущем изображении путем выполнения прогнозируемого кодирования в предыдущем изображении, что также дает возможность увеличить эффективность кодирования. Однако, даже когда применяется такое прогнозирование на основе компенсации движения, большое количество данных по сравнению с той частью изображения, где объект движется и появляется на заднем фоне, должно быть передано. Ввиду этого, не только компенсация движения в описанном выше направлении вперед, но также и компенсация движения в направлении назад или даже в обоих направлениях - вперед и назад - используется совместно, что дает возможность еще более увеличить эффективность кодирования.

Фактически, как показано на фиг. 8, в макроблоках данных кадров F0, F1, F2, F3 нулевого, первого, второго, третьего,... кадров видеосигнала подвижного изображения, подлежащего передаче, и в случае, когда имеют место изменения в изображении, что соответственно указано векторами движения x0, x1, x2, x3, ... последовательно между кадрами, устройство на передающей стороне обозначает кадры в заранее заданном количестве интервалов (например, исключая каждый следующий кадр), т.е. второй, четвертый,... как кадры интерполяции для так называемой заранее задаваемой кадровой интерполяции, как показано на фиг. 8, с целью формирования данных для передачи в виде интерполированных кадров F2X, F4X. . . . Далее, в отношении неинтерполированных кадров, устройство на передающей стороне выполняет заранее задаваемое колирование кадров F1, F2, F3,... для формирования подлежащих передаче данных - неинтерполированных кадров F1X, F3X,....

Например, разность SP2 (прогнозируемая ошибка) между данными кадров с компенсацией движения F3 и F2, разность SP4 между данными кадров с компенсацией движения F1 и F2 и разность между данными кадров, полученных в результате интерполяции данных кадров с компенсацией движения F1 и FT3, а также данные кадра F2 соответственно определяются в массивах макроблоков для сравнения разности (данных) SP1 данных кадра F2 и этих разностей. Затем данные, в составе которых

RU 2123769 C1

RU 2123769 C1

имеется минимальное количество данных, сформированных на основе SP1, SP4, выводятся на передачу как интерполированные данные F2X в массивах макроблоков. Аналогично интерполированные для передачи данные F4X, ... генерируются на основе соответствующих кадров интерполяции. Далее, данные неинтерполированных кадров F1, F3,... подвергаются обработке ДКП (дискретное косинусное преобразование) и неравномерному кодированию для формирования подлежащих передаче данных неинтерполированных кадров F1X, F3X, ...

Данные неинтерполированных кадров F1X, F3X,... и данные интерполированных кадров F2X, F4X, . . . передаются в устройство на приемной стороне совместно с векторами движения x0, x1, x3,...

С другой стороны, устройство на приемной стороне выполняет обработку декодированием, соответствующим обработке кодированием на предающей стороне, перед передачей данных (передаваемые данные неинтерполированных кадров F1X, F3X, . . . передаваемые интерполированные данные кадров F2X, F4X,... данные векторов движения x0, x1, x3) с тем, чтобы воспроизвести данные кадров F0, F1, F2, F3, . .. В результате компенсации движения выполняются не только в направлении вперед, но и в направлении назад или даже в обоих направлениях - вперед и назад, что дает возможность увеличить еще более эффективность кодирования.

Ниже приводится описание устройств кодирования и декодирования, имеющих описанные выше функции.

Устройство кодирования содержит, как показано на фиг. 9, схему 61 предварительной обработки, разделяющую видеосигнал ВС на сигнал яркости и цветоразностный сигнал, преобразовательные схемы 62а и 63б аналог/цифра (далее в тексте А/Ц схемы) для преобразования сигнала яркости и цветоразностного сигнала, соответственно поступающих со схемы 61, в цифровые сигналы, ЗУ кадровой памяти 63 для хранения данных яркости и цветоразностных данных (далее в тексте - данные изображения), поступающих со схем А-Ц 62а и 62б, схему 64 преобразования, схему 64 преобразования формата, служащую также для считывания данных из ЗУ 63 кадровой памяти в соответствии с блочным форматом, и кодирующее устройство (кодер) 65 для эффективного кодирования блока, поступающего со схемы 64 преобразования формата, в данные изображения.

В процессе работы схема 61 предварительной обработки разделяет входной видеосигнал ВС на сигналы яркости и цветоразностный. А/Ц схемы преобразования 62а и 62б соответственно преобразуют сигнал яркости и цветоразностный в данные яркости и цветоразностные данные группами по 8 бит каждая. ЗУ 63 хранит эти данные.

Схема 64 преобразования формата считывает в соответствии с блочным форматом, данные изображения (данные яркости и цветоразностные данные, хранящиеся в ЗУ 63 кадровой памяти). Кодер 65 кодирует данные изображения, считывая их таким образом с заданной эффективностью кодирования в выходной поток бит.

Этот поток бит направляется в устройство 80 декодирования изобретения через передающую среду 70, которой может быть либо трасса распространения сигнала, либо носитель записи изображения, например оптический диск, магнитный диск или магнитная лента и т.п.

Это устройство декодирования 80 содержит, как показано на фиг. 9, декодер 81, соответствующий кодеру 65, схему преобразования формата 82, служащую для преобразования данных изображения, воспроизведенных декодером 81, в формат кадра, ЗУ 83 кадровой памяти для хранения данных изображения, подаваемых со схемы 82 преобразования формата, Ц/А преобразовательные схемы 84а, 84б для преобразования данных яркости и цветоразностных данных, которые были считаны из ЗУ 83 кадровой памяти, в аналоговые сигналы, и послеобработочная схема 85 для смешения сигнала яркости и цветоразностного сигнала, подаваемых со схем 84а, 84б Ц/А преобразования, в результате чего формируется выходной видеосигнал.

Декодер 81 декодирует поток бит, выполняя операцию декодирования, соответствующую эффективному кодированию в кодере 65, для воспроизведения данных изображения в блочном формате. Схема 82 преобразует эти данные изображения и в формате кадра передает их на хранение в ЗУ 83 кадровой памяти.

Ц/А преобразовательные схемы 84а, 84б соответственно преобразуют данные яркости и цветоразностные данные, считанные из ЗУ 83 кадровой памяти, в сигнал яркости и в цветоразностный сигнал. Послеобработочная схема 81 смешивает эти два сигнала, формируя таким образом выходной видеосигнал.

Фактически послеобработочная схема 61 и Ц/А преобразовательные схемы 62а, 62б преобразуют сигнал яркости и цветоразностный сигнал в цифровой сигнал так же, как это делается и описано выше для сокращения количества данных, так что количество пикселей становится равным половине пикселей сигнала яркости в верхнем и нижнем направлениях, а также в левом и правом направлениях по отношению к сигналу яркости при дальнейшем осуществлении мультиплексирования по оси времени вдобавок к выдаче данных яркости и цветоразностных данных, полученных таким образом, в ЗУ 63 кадровой памяти.

Из ЗУ кадровой памяти 63 данные яркости и цветоразностные данные считываются в соответствии с блочным форматом, как описано выше, а именно, например, данные изображения одного кадра разделяются на N секций, как показано на фиг. 10А. Каждая секция формируется так, чтобы в ней содержалось М макроблоков, как показано на фиг. 10В. Каждый макроблок составляется из данных яркости в виде четырех блоков яркости Y1, Y2, Y3, Y4, каждый из которых состоит из 8 x 8 пикселей, в верхнем и нижнем, а также в левом и правом

RU 2123769 C1

RU 2123769 C1

направлениях, и из цветоразностных данных Сб, Ср цветных блоков, состоящих из 8 х 8 пикселей, соответствующих по размерам блокам яркости. Из ЗУ 63 кадровой памяти данные яркости и цветоразностные данные считываются так, что данные изображения затем располагаются последовательно в массивах макроблоков в пределах секций и последовательно в порядке Y1, Y2, Y3, Y4, Сб, Ср внутри макроблока. Данные изображения, считанные таким образом в соответствии с блочным форматом, затем пересылаются в кодер 65.

Кодер 65 содержит схему 101 определения вектора движения, как показано на фиг. 11. Эта схема 101 обнаруживает (в массивах макроблоков) вектор движения данных изображения, передаваемых дополнительно в соответствии с блочным форматом. Схема 101 обнаружения вектора движения обнаруживает в массивах макроблоков вектор движения текущего изображения как опорного на основе исходных изображений в направлении вперед/назад, хранящихся в ЗУ 63 кадровой памяти. Обнаружение осуществляется так, что минимальный вектор сумм абсолютных значений разностей между кадрами в массивах макроблоков становится соответствующим вектору движения. Определенный таким образом вектор движения выдается на схему 113 компенсации движения и далее, а внутрикадровые разности в массивах макроблоков выдаются на схему внутрикадрового прогнозирования вперед/назад/в обоих направлениях 103.

Эта схема 103 определяет прогнозируемый вид образцового блока с тем, чтобы на основе значения его параметра управлять схемой 104 прогнозируемого кодирования, выполняя переключения внутрикадрового прогнозирования вперед/назад/в обоих направлениях в массивах макроблоков. Схема 104 прогнозирования кодирования содержит схемы суммирования 104а, 104б, 104в и селективный переключатель 104г, эта схема в том случае, когда прогнозируемый режим кодирования оказывается внутрикадровым режимом, она отбирает разности между всеми пикселями (далее в тексте - разностные данные) входных данных изображения относительно соответствующих прогнозируемых изображений, а затем выдает отобранные данные на схему 105 ДКП (Дискретное Косинусное Преобразование).

Схема 105 выполняет ДКП-обработку данных входного изображения или разностных данных в массивах блоков, используя двумерную корреляцию видеосигнала для определения данных коэффициента и последующей передачи его на схему 106 квантования.

Схема квантования 106 квантует данные коэффициента, соблюдая размер шага квантования (шкала квантования), определяемый для каждого макроблока или секции с тем, чтобы затем выдать полученные таким образом данные на схему 107 неравномерного кодирования (с переменной длиной) (далее по тексту - НК) и на инверсную схему 108 квантования. При определении шага квантования необходимо следить за тем, чтобы величина этого шага не стала причиной разрыва упорядоченности в буферной памяти 109 передачи, которая будет описана ниже, вследствие возможного образования обратной связи по остаточным данным этого буфера 109. Значение этого шага квантования также передается на схему НК 107 и на инверсную схему квантования 108.

Схема НК 107 выполняет неравномерное кодирование квантованных данных, имеющих определенный шаг, в прогнозируемом режиме с учетом вектора движения, передавая затем все эти параметры в буфер 109 в качестве данных, подлежащих передаче.

Буфер 109 передачи временно хранит данные, подлежащие передаче с тем, чтобы в определенный момент считать их с заранее заданной скоростью чтения битов и упорядоченно передать их на выход в виде потока битов, а также передает сигнал управления квантованием в массивах макроблоков обратно на схему 106 квантования в соответствии с количеством остаточных данных, остающихся в буферной памяти, и для управления значением шага квантования. Таким образом, буфер передачи 109 регулирует количество данных, формируемых в виде потока битов для поддержания надлежащего остаточного количества данных в памяти (остаточная емкость) с тем, чтобы не допустить переполнения буфера и недостатка данных. Например, когда количество остаточных данных в буфере передачи 109 увеличивается до допустимого верхнего предела, буфер увеличивает шаг квантования схемы 108 с помощью управляющего сигнала, снижая таким образом количество квантуемых данных. С другой стороны, когда количество остаточных данных в буфере передачи 109 уменьшается до нижнего допустимого предела, буфер 109 передачи уменьшает с помощью сигнала управления квантованием размер шага квантования схемы квантования 106, увеличивая тем самым количество квантуемых данных.

Таким путем поток битов с выхода буфера 109 подается на блок 80 декодирования изображения через среду передачи 70, которая может представлять собой либо трассу распространения сигналов в эфире, либо носитель записанного изображения, например оптический диск, магнитный диск или магнитную ленту, передача производится с заранее заданной скоростью следования битов, что было описано выше.

С другой стороны, схема 108 инверсного квантования в инверсном режиме квантует уже квантованные данные, поступающие со схемы квантования 106 с тем, чтобы воспроизвести коэффициентные данные (при этом добавляются искажения, вызываемые квантованием), соответствующие выходным данным описанной выше схемы 105 ДКП, и выдать затем эти данные на схему 110 Инверсного Дискретного Косинусного Преобразования (далее по тексту - ИДКП). Схема 110 и ДКП выполняет ИДКП-обработку коэффициентных данных для воспроизведения данных изображения, соответствующих входным данным изображения во внутри-кадровом режиме кодирования, а также воспроизвести разностные данные, соответствующие выходным данным схемы 104 в режиме

RU 2123769 C1

RU 2123769 C1

прогнозирования вперед/назад/в обоих направлениях, передавая их затем на схему 111 суммирования.

Когда режим кодирования является прогнозирующим режимом вперед/назад/в обоих направлениях, на схему суммирования 111 подаются прогнозированные данные изображения с компенсацией движения со схемы 113 компенсации движения, что будет описано ниже, для суммирования прогнозированных данных изображения с компенсацией движения и разностных данных, тем самым воспроизводя данные изображения, соответствующие входным данным изображения.

Воспроизведенные таким образом данные изображения хранятся в кадровой памяти 112. Фактически, схема 108 инверсного квантования и схема суммирования 111 составляют схему местного декодирования, которая локально декодирует квантованные данные с выхода схемы 106 квантования с тем, чтобы затем записать декодированное изображение, получаемое таким образом, в кадровую память в качестве вперед или назад прогнозированного изображения. Общая кадровая память состоит из множества более мелких ячеек памяти. Здесь осуществляется переключение банка кадровой памяти. В соответствии с изображением, подлежащим кодированию, на выход выдается одиночный кадр в качестве вперед или назад прогнозированных данных изображения. Более того, в случае прогнозирования в обоих направлениях, вперед и назад прогнозируемые данные изображения, например, усредняются и выдаются на выход. Эти прогнозируемые данные изображения представляют полностью такие же изображения, как и изображения, воспроизведенные декодером 81, что будет описано ниже. Следующее изображение, подлежащее обработке, подвергается прогнозируемому кодированию вперед/назад/в обоих направлениях на основе этого прогнозированного изображения.

Так, данные изображения, считанные из кадровой памяти 112, направляются в схему 113 компенсации движения. Эта схема 113 выполняет компенсацию движения прогнозируемых данных изображения на основе вектора движения, а затем передает прогнозированные данные изображения с компенсацией движения на схему 104 прогнозируемого кодирования и на суммирующую схему 111.

Далее рассматривается декодер 81.

На вход декодера 81 поток битов подается через передающую среду 70, а затем этот поток подается на схему 202 Инверсного Кодирования с Переменной Длиной (ИКПД) через приемный буфер 201. Схема 202 ИКПД воспроизводит квантованные данные, вектор движения, прогнозируемый режим и шаг квантования из потока бит. Эти квантованные данные и размер шага квантования передаются в схему 203 инверсного квантования. Вектор движения подается на схему 207 компенсации движения, а режим прогнозирования устанавливается в суммирующей схеме 205.

Схема 203 инверсного квантования и суммирующая схема 205 совместно работают аналогично схеме локального декодирования кодера 61, а ЗУ кадровой памяти 206 и схема 207 компенсации движения соответственно выполняют те же функции, что и кадровая память 112 и схема 113 компенсации движения кодера 61. Декодирование осуществляется на основе квантованных данных, вектора движения, режима прогнозирования и шага квантования. В результате данные воспроизведения изображения выдаются с выхода схемы суммирования 205.

Как упоминалось выше, в обычных устройствах такого типа скорость кодирования потока бит, генерируемого кодером 65, устанавливается фиксированной в соответствии со скоростью передачи через передающую среду 70. При таком ограничении количество генерируемых данных, т.е. шаг квантования в схеме 106 кодера 65, находится под контролем. Другими словами, например, управление осуществлялось таким образом, что когда идет последовательность сложных изображений, шаг квантования задается увеличенным для сокращения количества генерируемых данных, тогда как при наличии последовательности более простых изображений шаг квантования уменьшается для увеличения генерируемых данных с тем, чтобы буфер 109 не переполнялся или его содержимое не было ниже установленного предела, благодаря чему поддерживается фиксированная скорость передачи.

Соответственно в обычных (известных) устройствах при наличии последовательности сложных изображений шаг квантования увеличивается, и при этом ухудшается качество изображения, тогда как, если имеет место последовательность простых изображений, то шаг квантования уменьшается. Как результат этого, одинакового качества изображения во всех возможных случаях получить не удается.

Кроме того, в случае регистрации потока бит на носителе регистрируемого изображения ограниченной емкости (по количеству данных) во избежание значительного ухудшения качества изображения относительно сложных изображений для всех случаев устанавливают высокую фиксированную скорость, которая уже не сказывается на качестве сложных изображений, но при этом уменьшается время регистрации.

Сущность изобретения.

Для решения изложенных выше проблем первый способ кодирования изображения согласно настоящему изобретению включает операции кодирования по меньшей мере части входного видеосигнала для генерации первых кодированных данных, определения скорости кодирования в каждый заранее заданный момент на основе количества данных в каждый заранее заданный момент существования первых кодированных данных и общего количества используемых данных и кодирования входного видеосигнала в каждый заранее заданный момент на основе скорости кодирования с тем, чтобы генерировать вторые кодированные данные.

Второй способ кодирования изображений согласно настоящему изобретению отличается тем, что в первом способе кодирования изображения по меньшей мере часть входного видеосигнала квантуется с фиксированным шагом квантования для

RU 2123769 C1

RU 2123769 C1

генерации первых кодированных данных.

Третий способ кодирования изображений согласно настоящему изобретению отличается тем, что включает операции заранее определяемого прогнозируемого кодирования или заранее задаваемого преобразующего кодирования по меньшей мере части входного видеосигнала для генерации первых коэффициентов данных, квантования этих первых коэффициентов данных с фиксированным шагом для генерации (выработки) первых квантованных данных, что позволяет подвергнуть первые квантованные данные неравномерному кодированию (с переменной длиной) для формирования первого потока бит, определения скорости кодирования в каждый заранее заданный момент времени на основе количества данных в первом потоке бит и общего количества пригодных к использованию данных, выполнения заранее задаваемого прогнозирующего кодирования и/или заранее задаваемого преобразующего кодирования входного видеосигнала для генерации вторых коэффициентов данных, квантования этих вторых коэффициентов данных с шагом квантования, выбираемым на основе скорости кодирования в каждый заранее заданный момент времени, для генерации вторых квантованных данных и разрешения вторым квантованным данным подвергнуться неравномерному кодированию (с переменной длиной) для формирования второго потока бит.

Третий способ кодирования изображений согласно настоящему изобретению отличается тем, что в первом способе кодирования изображений общее количество пригодных для использования данных пропорционально распределяется в каждый заранее заданный момент времени, благодаря чему определяется скорость кодирования в каждый заранее заданный момент времени.

Пятый способ кодирования изображения согласно настоящему изобретению отличается тем, что при использовании четвертого способа кодирования изображений скорость кодирования определяется в каждом кадре на основе количества данных в каждом кадре первого потока бит и общего количества пригодных к использованию данных.

Шестой способ кодирования изображений согласно настоящему изобретению отличается тем, что при осуществлении четвертого способа кодирования изображений скорость кодирования определяется в каждой Группе Изображений (GOP) на основе количества данных из состава по меньшей мере части каждой GOP, состоящей из множества кадров из первого потока бит и общего количества пригодных к использованию данных.

Седьмой способ кодирования изображений согласно настоящему изобретению отличается тем, что при использовании шестого способа скорость кодирования в каждой ГИ определяется на основе количества данных, относящихся к изображению, закодированному внутри кадра, и прогнозируемому вперед кодированному изображению из состава ГИ.

Восьмой способ кодирования изображений согласно настоящему изобретению отличается тем, что при использовании четвертого способа кодирования общее количество пригодных данных пропорционально распределяется в зависимости от количества данных в первом потоке бит в каждый отрезок времени, на основе чего скорость кодирования определяется в каждый заданный отрезок времени.

Девятый способ кодирования изображений согласно настоящему изобретению включает операции определения предела кодирования в каждом заранее заданном массиве изображения во входном видеосигнале, установки скорости кодирования в каждом заранее заданном массиве изображения на основе предела кодирования и общего количества пригодных данных и выполнения кодирования входного видеосигнала таким образом, что скорости кодирования соответственных массивов изображения соответствуют установленной скорости кодирования для каждого массива изображения.

Десятый способ кодирования изображений согласно настоящему изобретению отличается тем, что при использовании девятого способа заранее задаваемым массивом изображения является кадр.

Одиннадцатый способ кодирования изображений согласно настоящему изобретению отличается тем, что при осуществлении девятого способа кодирования заранее заданным массивом изображения выбирается ГИ, состоящая из множества кадров.

Двенадцатый способ кодирования изображений согласно настоящему изобретению отличается тем, что при осуществлении девятого способа кодирования изображения выполняется заранее задаваемое прогнозирующее кодирование и/или заранее задаваемое преобразовательное кодирование по меньшей мере части входного видеосигнала для формирования коэффициентных данных и квантования этих данных с фиксированным шагом для определения предела кодирования.

Первое устройство для кодирования изображений согласно настоящему изобретению содержит первое кодирующее средство для кодирования по меньшей мере части входного видеосигнала и последующего генерирования первых кодированных данных, средство управления кодированием для определения скорости кодирования в каждый заранее задаваемый отрезок времени на основе наличия в каждый заранее задаваемый момент количества первых кодированных данных, поступающих от первого кодирующего средства, и общего количества пригодных к использованию данных и вторые кодированные данные для кодирования входного видеосигнала в каждый заранее задаваемый отрезок времени на основе скорости кодирования, заранее задаваемой в каждый отрезок времени средством управления кодированием для генерирования вторых кодированных данных.

Второе устройство кодирования изображения согласно настоящему изобретению отличается тем, что в первом устройстве кодирования изображения первое кодирующее средство содержит средство

RU 2123769 C1

RU 2123769 C1

квантования, служащее для квантования по меньшей мере части входного видеосигнала с фиксированным шагом.

Третье устройство кодирования изображений согласно настоящему изображению отличается тем, что при используемом первом устройстве кодирования изображения средство управления кодированием осуществляет пропорциональное распределение общего количества пригодных данных в зависимости от наличия определенного количества данных в каждый заранее заданный момент (отрезок) времени с тем, чтобы определять скорость кодирования в каждый заранее задаваемый отрезок времени.

Четвертое устройство кодирования изображений согласно настоящему изобретению содержит средство кодирования, служащее для выполнения заранее задаваемого прогнозирующего кодирования и/или заранее задаваемого преобразовательного кодирования по меньшей мере части входного видеосигнала для генерирования коэффициентных данных, первое квантующее средство для квантования первых коэффициентных данных, поступающих с первого кодирующего средства, с фиксированным шагом квантования для вырабатывания первых квантованных данных, первое средство неравномерного кодирования (с переменной длиной) для обеспечения возможности для квантованных данных, приходящих с первого средства квантования, подвергнуться неравномерному кодированию (с переменной длиной) для формирования первого потока бит, средство управления кодированием для определения скорости кодирования в каждый заранее задаваемый отрезок времени на основе количества данных в первом потоке бит, поступающем с первого средства неравномерного кодирования (с переменной длиной), и общего количества пригодных для использования данных, второе кодирующее средство, служащее для выполнения заранее задаваемого прогнозирующего кодирования и/или прогнозирующего преобразовательного кодирования входного видеосигнала для выработки вторых коэффициентных данных, второе средство квантования, квантующее вторые коэффициентные данные с выхода второго средства кодирования с шагом квантования, задаваемым на основе скорости кодирования в каждый заранее задаваемый отрезок времени средством управления кодированием для выработки вторых квантованных данных, и второе средство неравномерного кодирования (с переменной длиной), позволяющее вторым квантованным данным с выхода квантующего средства пройти кодирование с переменной длиной для формирования второго потока бит.

Пятое устройство кодирования изображения согласно настоящему изобретению отличается тем, что при использовании четвертого устройства кодирования средство управления кодированием определяет скорость кодирования в каждом кадре на основе количества данных в каждом кадре первого потока бит и общего количества пригодных данных.

Шестое устройство кодирования изображений согласно настоящему изобретению отличается тем, что при использовании четвертого устройства кодирования средство управления кодированием определяет скорость кодирования в каждой ГИ на основе количества данных по крайней мере в части каждой ГИ, состоящей из множества кадров в первом потоке бит, и в общем количестве пригодных данных.

Седьмое устройство кодирования изображений согласно настоящему изобретению отличается тем, что при использовании шестого устройства кодирования изображений средство управления кодированием определяет скорость кодирования в каждой ГИ на основе количества данных в закодированном внутри кадра изображении и прогнозированно закодированном вперед изображении в ГИ.

Восьмое устройство кодирования изображений согласно настоящему изобретению средство управления кодированием пропорционально распределяет общее количество пригодных данных в зависимости от количества данных в первом потоке бит в каждый заранее задаваемый отрезок времени, определяя таким образом скорость кодирования в каждый задаваемый момент времени.

Девятое устройство кодирования изображения согласно настоящему изобретению содержит средство вычисления предела, служащее для расчета (определения) предела кодирования в каждом заранее задаваемом массиве изображения входного видеосигнала, средство установки скорости кодирования, служащее для установки скорости кодирования в каждом заранее задаваемом массиве изображения на основе показателей предела кодирования, поступающих со средства вычисления предела и общего количества пригодных к использованию данных, и кодирующее средство, дающее возможность входному видеосигналу пройти кодирование таким образом, что скорости кодирования соответствующих массивов изображения находятся в соответствии со скоростью кодирования каждого массива изображения, установленной средством установки скорости кодирования.

Десятое устройство кодирования изображения согласно настоящему изобретению отличается тем, что при использовании девятого устройства кодирования изображения средство вычисления помех определяет (предел) кодирования для каждого кадра.

Одиннадцатое устройство для кодирования изображений согласно настоящему изобретению отличается тем, что средство вычисления определяет (предел) кодированию для каждой ГИ, состоящей из множества кадров.

Двенадцатое устройство кодирования изображения согласно настоящему изобретению отличается тем, что при использовании девятого устройства кодирования изображения средство вычисления предела выполняет заранее задаваемое прогнозирующее кодирование и/или заранее задаваемое преобразовательное кодирование по меньшей мере части входного видеосигнала для вырабатывания коэффициентных данных

RU 2123769 C1

RU 2123769 C1

и последующего квантования этих данных при фиксированном шаге квантования, на основе чего вычисляется (определяется) предел кодирования.

Первый носитель информации для регистрации изображения согласно настоящему изобретению отличается тем, что регистрируется (записывается) второй поток бит, получаемый путем кодирования по крайней мере части входного видеосигнала для генерации первых кодированных данных, по которым определяют скорость кодирования в каждый заранее задаваемый отрезок времени на основе наличия в каждый заранее задаваемый отрезок времени количества первых кодированных данных и общего количества пригодных для использования данных с тем, чтобы кодировать входной видеосигнал в каждый заранее задаваемый отрезок времени на основе кодированных данных.

Второй носитель информации для регистрации (записи) изображения согласно настоящему изобретению отличается тем, что регистрируется второй поток бит, получаемых путем выполнения заранее задаваемого прогнозирующего кодирования и/или заранее задаваемого преобразовательного кодирования по крайней мере части входного видеосигнала для выработки первых коэффициентных данных и последующего квантования этих данных с фиксированным шагом квантования с тем, чтобы получить первые квантованные данные, которые получат возможность подвергнуться неравномерному кодированию (с переменной длиной) для формирования первого потока бит, по которому определяется скорость кодирования в каждый заранее задаваемый отрезок времени на основе количества данных в первом потоке бит и общего количества пригодных к использованию данных с тем, чтобы затем выполнить заранее задаваемое прогнозирующее кодирование и/или заранее задаваемое преобразовательное кодирование входного видеосигнала для выработки вторых коэффициентных данных и их последующего квантования с фиксированным шагом, задаваемым на основе скорости кодирования в каждый заранее задаваемый отрезок времени, и последующей выработки вторых квантованных данных, которые получат возможность пройти неравномерное кодирование (с переменной длиной).

Третий носитель информации для регистрации (записи) изображения согласно настоящему изобретению отличается тем, что на регистрацию выводятся кодирование данные, полученные при определении предела кодирования в каждом заранее задаваемом массиве изображения во входном видеосигнале с тем, чтобы установить скорость кодирования в каждом заранее задаваемом массиве изображения на основе показателей предела кодирования и количества пригодных к использованию данных и выполнить кодирование входного видеосигнала таким образом, чтобы скорости кодирования соответствующих массивов изображения находились в соответствии с установленной для каждого массива изображения скоростью кодирования.

В соответствии с первым способом кодирования изображения согласно настоящему изобретению скорость кодирования в каждый заранее заданный отрезок времени определяется на основе количества данных в каждый заранее задаваемый отрезок времени из числа первых кодированных данных, получаемых путем кодирования по крайней мере части входного видеосигнала, и общего количества пригодных данных для дальнейшего кодирования входного видеосигнала в каждый заранее задаваемый отрезок времени на основе определенной скорости кодирования и для последующего вырабатывания вторых кодированных данных.

В соответствии с третьим способом кодирования изображений согласно настоящему изобретению при осуществлении первого способа кодирования изображения общее количество пригодных данных пропорционально распределяется в зависимости от количества данных в каждый заранее задаваемый отрезок времени для определения скорости кодирования в каждый заранее заданный отрезок времени для дальнейшего кодирования входного видеосигнала в каждый заданный отрезок времени на основе определенной скорости кодирования и для вырабатывания вторых кодированных данных.

В соответствии с четвертым способом кодирования изображения согласно настоящему изобретению сначала по крайней мере только часть входного сигнала подвергается последовательно заранее задаваемому прогнозирующему кодированию и/или заранее задаваемому преобразовательному кодированию, обработке в виде квантования и неравномерному кодированию (с переменной длиной) при фиксированном шаге квантования для выработки первого потока бит и определения скорости кодирования на основе количества данных в первом потоке бит и общего количества пригодных данных. Затем для формирования второго потока бит весь входной видеосигнал подвергается обработке в виде заранее задаваемого прогнозирующего кодирования и/или заранее задаваемого преобразовательного кодирования, квантования и кодированию с переменной длиной при фиксированном шаге квантования, устанавливаемого на основе скорости кодирования в каждый заранее задаваемый отрезок времени.

В соответствии с пятым способом кодирования изображений согласно настоящему изобретению при осуществлении четвертого способа кодирования изображений согласно настоящему изобретению скорость кодирования определяется в каждом кадре на основе количества данных в одном кадре потока бит и общего количества пригодных данных. Затем для формирования второго потока бит входной видеосигнал подвергается заранее задаваемому прогнозирующему кодированию и/или заранее задаваемому преобразовательному кодированию, а также обработке квантованием и неравномерным кодированием (с переменной длиной) при фиксированном шаге квантования, задаваемом на основе скорости кодирования в каждом кадре.

В соответствии с шестым способом кодирования изображений согласно настоящему изобретению берут за основу

RU 2123769 C1

RU 2123769 C1

четвертый способ кодирования изображения и определяют скорость кодирования в каждой ГИ на основе количества данных по меньшей мере в части каждой ГИ, состоящей из множества кадров первого потока бит, и общего количества пригодных данных.

В соответствии с седьмым способом кодирования изображений согласно настоящему изобретению за основу берут шестой способ кодирования изображения и определяют скорость кодирования в каждой ГИ на основе количества данных относительно закодированного внутри кадра изображения и прогнозированного вперед изображения в составе ГИ. Затем входной видеосигнал подвергается обработке в виде (для формирования второго потока бит) заранее задаваемого прогнозирующего кодирования и/или заранее задаваемого преобразовательного кодирования, квантизации с фиксированным шагом, задаваемым на основе скорости кодирования в каждой ГИ, а также обработке кодированием с переменной длиной.

В соответствии с восьмым способом кодирования изображений согласно настоящему изобретению за основу берут четвертый способ кодирования изображений и общее количество пригодных к использованию данных пропорционально распределяют в зависимости от количества данных в первом потоке бит в каждый заранее задаваемый отрезок времени для определения скорости кодирования в каждый заранее задаваемый отрезок времени. Затем для формирования второго потока бит входной видеосигнал подвергается обработке в виде заранее задаваемого прогнозирующего кодирования и/или заранее задаваемого преобразовательного кодирования, квантизации с фиксированным шагом квантизации, задаваемым на основе скорости кодирования в каждый заранее задаваемый отрезок времени и кодирования с переменной длиной.

В соответствии с девятым способом кодирования изображений согласно настоящему изобретению предел кодирования в каждом заранее задаваемом массиве изображения во входном видеосигнале определяют для установления скорости кодирования в каждом заранее задаваемом массиве изображения на основе показателей предела кодирования и общего количества пригодных к использованию данных. Затем входной видеосигнал кодируется таким образом, что скорости кодирования в соответствующих массивах изображения находятся в соответствии с установленной скоростью кодирования для каждого массива изображения.

В соответствии с десятым способом кодирования изображения согласно настоящему изобретению за основу берут девятый способ кодирования изображений и определяют помехи кодированию для каждого кадра входного видеосигнала для определения скорости кодирования в каждом кадре. Затем входной видеосигнал кодируется так, что скорости кодирования соответствующих кадров соответствуют установленной скорости кодирования в каждом кадре.

В соответствии с одиннадцатым способом кодирования изображений за основу снова берут девятый способ и определяют порог кодирования в каждой ГИ входного видеосигнала для определения скорости кодирования в каждой ГИ. Затем входной видеосигнал кодируется так, что скорости кодирования в Группах Изображений находятся в соответствии с установленной для каждой ГИ скоростью кодирования.

В соответствии с двенадцатым способом кодирования изображений согласно настоящему изобретению также за основу берут девятый способ кодирования изображений, и коэффициентные данные, получаемые путем заранее задаваемого прогнозирующего кодирования и/или заранее задаваемого преобразовательного кодирования по меньшей мере части входного видеосигнала, квантуют с фиксированным шагом для определения таким образом предельных показателей. Затем входной видеосигнал кодируется так, что скорости кодирования соответствующих массивов изображений поддерживаются в соответствии со скоростью кодирования каждого массива изображения.

Далее, в соответствии с первым устройством кодирования изображений согласно настоящему изобретению скорость кодирования в каждый заранее задаваемый отрезок времени определяется на основе количества данных в каждый заданный отрезок времени в составе первых кодированных данных, полученных путем кодирования по меньшей мере части входного видеосигнала в каждый заранее задаваемый отрезок времени, и общего количества пригодных данных для дальнейшего кодирования всего входного видеосигнала в каждый заранее задаваемый отрезок времени на основе определенной скорости кодирования и последующей выработки вторых кодированных данных.

В соответствии с вторым устройством кодирования изображений согласно настоящему изобретению за основу берут первое устройство и по меньшей мере часть входного видеосигнала квантуется при фиксированном шаге квантования для выработки таким образом первых кодированных данных, по которым определяют скорость кодирования, с которой кодируется входной видеосигнал в каждый заранее задаваемый отрезок времени, в результате чего вырабатываются вторые кодированные данные.

В соответствии с третьим устройством кодирования изображений согласно настоящему изобретению используется первое устройство кодирования изображения, в котором общее количество пригодных для использования данных пропорционально распределяется в зависимости от количества данных в каждый заранее задаваемый отрезок времени для определения скорости кодирования в каждый заранее задаваемый момент времени и затем с этой скоростью кодируется входной видеосигнал в каждый заданный момент, в результате чего вырабатываются вторые кодированные данные.

В соответствии с четвертым устройством кодирования изображения согласно настоящему изобретению по меньшей мере часть входного видеосигнала подвергается обработке в виде заранее заданного

RU 2123769 C1

RU 2123769 C1

прогнозирующего кодирования и/или заранее заданного преобразовательного кодирования, а также квантования с фиксированным шагом и неравномерного кодирования (с переменной длиной) для формирования первого потока бит с тем, чтобы определить скорость в каждый заранее задаваемый отрезок времени на основе количества данных в первом потоке бит и общего количества данных, пригодных к использованию. Затем заранее задаваемому прогнозируемому кодированию и/или преобразовательному кодированию, а также квантованию и кодированию с переменной длиной с шагом квантования, задаваемым на основе скорости кодирования в каждый задаваемый отрезок времени, подвергается весь входной видеосигнал, в результате чего формируется второй поток бит.

В соответствии с пятым устройством кодирования изображения согласно настоящему изобретению во взятом за основу четвертом устройстве кодирования изображения скорость кодирования определяется в каждом кадре на основе количества данных в каждом кадре первого потока бит и общего количества пригодных данных. Затем обработке в виде заранее задаваемого прогнозирующего кодирования и/или прогнозирующего преобразовательного кодирования, а также квантования с шагом квантования, задаваемым на основе скорости кодирования в каждом кадре, и кодированию с переменной длиной подвергается весь входной видеосигнал, в результате чего формируется второй поток бит.

В соответствии с шестым устройством кодирования изображений согласно настоящему изобретению в принятом за основу четвертом устройстве кодирования изображений скорость кодирования определяется в каждой ГИ на основе количества данных по меньшей мере в части каждой ГИ, состоящей из множества кадров в первом потоке бит, и общего количества пригодных данных. Затем обработке подвергается весь входной видеосигнал, а именно заданному прогнозирующему кодированию и/или заданному преобразовательному кодированию, квантованию с шагом квантования, задаваемым на основе скорости кодирования в каждой ГИ, и кодированию с переменной длиной, в результате чего формируется второй поток бит.

В соответствии с седьмым устройством кодирования изображений согласно настоящему изобретению в принятом за основу шестом устройстве кодирования изображения скорость кодирования в каждой ГИ определяется на основе количества данных в изображениях, кодированных внутри кадра, и прогнозируемых вперед кодированных изображениях в составе ГИ. Затем входной видеосигнал подвергается обработке в виде заранее задаваемого прогнозирующего кодирования и/или заранее задаваемого преобразовательного кодирования, квантования с шагом, определяемым скоростью кодирования в каждой ГИ, и кодирования с переменной длиной, в результате чего формируется второй поток данных.

В соответствии с восьмым устройством кодирования изображения согласно настоящему изобретению в принимаемом за основу четвертом устройстве кодирования изображения общее количество пригодных к использованию данных пропорционально распределяется в зависимости от количества данных в первом потоке бит в каждый заранее задаваемый отрезок времени для определения скорости кодирования в каждый задаваемый момент. Затем входной видеосигнал подвергается обработке в виде заранее задаваемого прогнозирующего кодирования и/или заранее задаваемого преобразовательного кодирования, квантования с шагом, определяемым на основании скорости кодирования в каждый заранее заданный момент, и неравномерному кодированию (с переменной длиной), в результате чего формируется второй поток данных.

В соответствии с девятым устройством кодирования изображения согласно настоящему изобретению предел кодирования определяется в каждом задаваемом массиве изображения входного видеосигнала для установления скорости кодирования в каждом заданном массиве изображения на основе параметров предела кодирования и общего количества данных, пригодных к использованию. Затем входной видеосигнал кодируется так, что скорости кодирования соответственных массивов изображения поддерживаются в соответствии с установленной скоростью кодирования в каждом массиве изображения.

В соответствии с девятым устройством кодирования изображения согласно настоящему изобретению в принимаемом за основу девятом устройстве кодирования порог кодирования определяется в каждом кадре входного видеосигнала для определения скорости кодирования в каждом кадре. Затем входной видеосигнал кодируется так, что скорости кодирования в соответственных кадрах находятся в соответствии с установленной скоростью кодирования в каждом кадре.

В соответствии с одиннадцатым устройством кодирования изображения согласно настоящему изобретению за основу берут снова девятое устройство кодирования и помехи кодированию определяют в каждой ГИ входного видеосигнала для определения скорости кодирования в каждой GOP. Затем входной видеосигнал кодируется так, что скорости кодирования соответственных ГИ поддерживаются в соответствии со скоростью кодирования, установленной для каждой ГИ.

В соответствии с двенадцатым устройством кодирования изображений согласно настоящему изобретению за основу также берут девятое устройство кодирования, и коэффициентные данные, получаемые путем выполнения заранее задаваемого прогнозирующего кодирования и/или заранее задаваемого преобразовательного кодирования по меньшей мере части входного видеосигнала, квантуют с фиксированным шагом для определения предела кодирования. Затем входной видеосигнал кодируется таким образом, что скорости кодирования соответственных массивов изображения находятся в соответствии со скоростью кодирования, установленной для каждого массива изображения.

Кроме того, в соответствии с первым носителем для регистрации (записи)

RU 2123769 C1

RU 2123769 C1

изображения согласно настоящему изобретению на регистрацию выводится второй поток бит, получаемых путем определения скорости кодирования в каждый заранее задаваемый отрезок времени на основе количества данных в каждый момент или отрезок времени из состава первых кодированных данных, получаемых кодированием по меньшей мере части входного видеосигнала, и общего количества пригодных к использованию данных с тем, чтобы кодировать входной видеосигнал в каждый заранее задаваемый момент времени на основе определенной скорости кодирования.

В соответствии с вторым носителем информации для регистрации изображения согласно настоящему изобретению регистрируется второй поток бит, получаемый путем выполнения заранее задаваемого прогнозирующего кодирования и/или заранее задаваемого преобразовательного кодирования по меньшей мере части входного видеосигнала, а также квантования и неравномерного кодирования (с переменной длиной) с фиксированным шагом для генерации первого потока бит с целью определения скорости кодирования в каждый заданный отрезок времени на основе количества данных в первом потоке бит и общего количества пригодных данных с тем, чтобы подвергнуть обработке входной видеосигнал, а именно заранее задаваемому прогнозирующему кодированию и/или заранее задаваемому преобразовательному кодированию, а также квантованию и кодированию с переменной длиной с шагом квантования, определяемым на основе скорости кодирования в каждый заранее задаваемый отрезок.

В соответствии с третьим носителем информации для регистрации изображения согласно настоящему изобретению на регистрацию выводятся кодированные данные, получаемые при определении предела кодирования в каждом заданном массиве изображения входного видеосигнала с тем, чтобы установить скорость кодирования каждого заранее задаваемого массива изображения на основе параметров предела кодирования и общего количества данных, пригодных для использования, чтобы кодировать входной видеосигнал таким образом, чтобы скорости кодирования соответственных массивов изображения были в соответствии со скоростью кодирования каждого массива изображения.

Краткое описание фигур.

Фиг. 1 - блок-схема, показывающая связи основных частей устройства кодирования изображений.

Фиг. 2 - блок-схема алгоритма, поясняющего работу первой кодирующей схемы, входящей в состав устройства кодирования по фиг. 1.

Фиг. 3 - блок-схема алгоритма, поясняющего работу второй схемы кодирования, входящей в состав устройства кодирования изображения по фиг. 1.

Фиг. 4 - относительное расположение изображений в составе группы изображений GOP.

Фиг. 5 - относительное расположение изображений и сигнал управления кодирования в каждой группе изображений GOP.

Фиг. 6 - блок-схема алгоритма, также поясняющего работу схемы кодирования, входящей в состав устройства кодирования изображений по фиг. 1.

Фиг. 7 - вид изображения для пояснения принципа прогнозирующего кодирования.

Фиг. 8 - диаграмма изображений, поясняющая принцип прогнозирующего кодирования с компенсацией движения.

Фиг. 9 - блок-схема устройства (блока) кодирования изображений и устройства (блока) декодирования изображений согласно настоящему изобретению.

Фиг. 10 - конфигурация кадра, макро-блока и секции.

Фиг. 11 - блок-схема обычного (известного) кодера.

Фиг. 12 - блок-схема обычного (известного) декодера.

Наилучший вариант реализации настоящего изобретения

Далее приводится описание со ссылкой на прилагаемые чертежи одного из вариантов способа кодирования изображений, устройства кодирования изображений и носителя информации для регистрации (записи) изображений согласно настоящему изобретению.

Устройство кодирования изображений согласно настоящему изобретению содержит, как показано на фиг. 10, например, первую схему кодирования 10 для кодирования входного видеосигнала и выработки первых кодированных данных, схему кодирования 30 изображений для определения скорости кодирования в каждый заранее задаваемый отрезок времени на основе количества кодированных данных в каждый заранее задаваемый момент, поступающих с первой кодирующей схемы 10, и общего количества пригодных к использованию данных и вторую схему кодирования 40 для кодирования входного видеосигнала в каждый заранее задаваемый отрезок времени на основе скорости кодирования, определяемой схемой управления кодированием 30 и последующей выработки вторых кодированных данных.

Более конкретно, первая схема 10 кодирования содержит, как показано на фиг. 1, ЗУ 12 кадровой памяти для хранения входных данных изображений, представляющих собой входной видеосигнал, схему 11 определения вектора движения, служащую для обнаружения этого вектора во входных данных изображений, хранящихся в ЗУ 12 кадровой памяти, ЗУ 22 кадровой памяти, хранящее прогнозированные данные изображений, схему 23 компенсации движения, служащую для выполнения компенсации движения прогнозированных данных изображений, которые были считаны из ЗУ 22 кадровой памяти на основе данных о векторе движения, поступающих со схемы 11 обнаружения вектора движения, схему 14 прогнозирующего кодирования для кодирования входных данных изображения на основе прогнозированных данных изображений с компенсацией движения, поступающих со схемы 23 компенсации движения, схему 15 ДКП, выполняющую кодирование разностей в виде Дискретного Косинусного Преобразования (далее по тексту - ДКП) разностей и т.п., которые называют прогнозируемыми ошибками,

RU 2123769 C1

RU 2123769 C1

поступающих со схемы 14 прогнозирующего кодирования, схему 16 квантования для квантования коэффициентных данных со схемы 15 ДКП, с фиксированным шагом квантования для выработки квантованных данных, схему 17 неравномерного кодирования (с переменной длиной) (далее по тексту - НК), позволяющую квантованным данным со схемы 16 подвергнуться обработке НК с тем, чтобы выдать на выход кодированные с переменной длиной данные, схему 18 инверсного квантования данных, поступающих со схемы 16 квантования для воспроизведения коэффициентных данных, схему 20 Инверсного Дискретного Косинусного Преобразования (далее по тексту ИДКП), служащую для выполнения декодирования, например, в виде ИДКП обработки коэффициентных данных со схемы 18 инверсного квантования, и схему 21 суммирования для суммирования данных с выхода схемы 20 ИДКП и прогнозируемых данных изображения с компенсацией движения с выхода схемы 23 компенсации движения, в результате чего вырабатываются прогнозируемые данные, относящиеся к данным следующего изображения, передаваемых затем в ЗУ кадровой памяти 22.

Далее, вторая схема кодирования 40 содержит, как показано на фиг. 1, элемент задержки 43 для задержки входных данных изображения, ЗУ 52 кадровой памяти для хранения прогнозируемых данных изображения, схему 53 компенсации движения прогнозируемых данных изображения, которые были считаны из ЗУ 52 кадровой памяти с учетом вектора движения со схемы 11 обнаружения этого вектора, схему 44 прогнозирующего кодирования входных данных изображений, задержанных элементом задержки 43, на основе прогнозируемых данных изображений с компенсацией движения, поступающих со схемы 53 компенсации движения, схему 45 ДКП, выполняющую кодирование, например ДКП-обработку разностей и т.п., поступающих со схемы 44 прогнозирующего кодирования, для выработки коэффициентных данных, схему 33 установки шкалы квантования (шага квантования) на основе скорости кодирования, определяемой схемой 30 управления кодированием, схему 46 квантования коэффициентных данных, поступающих со схемы 45 ДКП, с шагом квантования, задаваемым схемой 33 установки шкалы квантования, схему 47 НК, дающую возможность подвергнуть квантованные данные со схемы 46 НК-обработке и вывести после этого данные на выход, буфер передачи 49 для временного хранения данных с выхода схемы 47 НК для последующего вывода их с фиксированной скоростью на передачу, схему 48 для инверсного квантования данных со схемы квантования 46 и дальнейшего воспроизведения коэффициентных данных, схему ИДКП 50, выполняющую декодирование, например, в виде ИДКП обработки коэффициентных данных с выхода схемы 48 инверсного квантования с целью воспроизведения разности, и схему 51 суммирования разности с выхода схемы 50 ИДКП и прогнозируемых данных с компенсацией движения с выхода схемы 53 компенсации движения с тем, чтобы выработать прогнозируемые данные изображений, относящиеся к следующим входным данным изображений, и передать эти прогнозируемые данные в ЗУ 52 кадровой памяти.

В настоящем варианте устройства кодирования обработке, которую выполняет первое средство кодирования 10, а именно прогнозирующему кодированию, ДКП-обработке, квантованию с фиксированным шагом и кодированию (с переменной длиной) подвергается входной видеосигнал, скорость кодирования битов определяется средством управления кодированием 30 на основе количества кодированных с переменной длиной данных, которые представляют собой первый получаемый поток бит, и емкости (по данным) носителя 55 записанного изображения, например оптического диска, магнитного диска или магнитной ленты и т.п., или же общего количества пригодных к использованию данных, определяемого скоростью передачи бит через передающую среду. Соответственно второе средство 40 кодирования выполняет прогнозирующее кодирование, ДКП-обработку, квантование и неравномерное кодирование (с переменной длиной) входных данных изображения во второй раз.

Как показано на фиг. 2, в настоящем варианте реализации изобретения в шаге алгоритма 1 схема квантования 16 из состава первого средства кодирования 10 устанавливает шаг квантования равным, например, 1 для квантования коэффициентных данных, переданных со схемы 15 ДКП для вырабатывания квантованных данных. Счетчик 31 средства управления кодированием 30 подсчитывает количество данных, кодированных с переменной длиной (первый поток данных), получаемых в случае представления возможности кодирования (с переменной длиной) квантованных данных в заранее задаваемый отрезок времени, например за время одного кадра, для определения количества формируемых в каждом кадре количества кодов, указывающих предел (порог) кодирования.

В шаге 2 алгоритма схема 32 вычисления скорости передачи битов определяет количество кодов, предназначенных для каждого кадра на основе предела или порога (количества формируемых кодов) каждого кадра и общего количества пригодных данных.

В шаге 3 алгоритма схема квантования 46, второго средства кодирования 40 квантует данные, приходящие со схемы 45 ДКП, с шагом квантования, задаваемым на основе количества кодов, предназначенных для вырабатывания квантованных данных.

Подаваемые на вход данные изображений временно хранятся в ЗУ 12 кадровой памяти, откуда они затем считываются в соответствии с блочным форматом, что описано в известных технических решениях.

Схема 11 обнаружения вектора движения считывает необходимые данные изображения из ЗУ 12 кадровой памяти массивами макроблоков для обнаружения вектора движения, что было описано выше. Эта схема 11 обнаруживает в массивах макроблоков

RU 2123769 C1

RU 2123769 C1

вектор движения текущего (опорного) изображения, используя расположенное впереди или позади исходное изображение из ЗУ 12 кадровой памяти. Здесь обнаружение вектора движения осуществляется так, что вектор движения, для которого, например, сумма абсолютных значений разностей между кадрами в массивах макроблоков становится минимальной, признается соответствующим искомому вектору движения. Обнаруженный вектор движения передается на схемы 23, 53 и т.п. компенсации движения, а сумма абсолютных значений разностей между кадрами в массивах макроблоков подается на схему 13 прогнозирования внутри кадра/вперед/назад/ в обоих направлениях (схема принятия решения).

Эта схема 13 определяет режим прогнозирования на основе упомянутого выше значения, используя его для управления схемой 14 прогнозирующего кодирования, которая переключает режимы прогнозирования "внутрикадровый/вперед/назад/в обоих направлениях" в массивах блоков.

Прогнозирующая кодирующая схема 14 содержит, как можно видеть из фиг. 1, схемы суммирования 14а, 14б, 14в и селекторный (смены режима) переключатель 14г. Когда работа идет во внутрикадровом режиме, этот переключатель манипулирует непосредственно входными данными изображения, но если же задан режим прогнозирования "вперед/назад/ в обоих направлениях", этот переключатель манипулирует разностями (далее по тексту - разностные данные) в каждом из пикселей входных данных изображения, относящихся к соответственным прогнозируемым изображениям. С селекторного переключателя 14 выбранные им данные подаются на схему 15 ДКП.

Схема 15 ДКП выполняет (в масштабе массивов блоков) ДКП-обработку входных данных изображений или разностных данных, приходящих с селекторного переключателя 14г, с использованием двумерной корреляции видеосигнала, а затем получаемые таким образом коэффициентные данные подаются на схему 16 квантования.

Схема 16 квантования квантует коэффициентные данные, поступающие со схемы 15 ДКП, с шагом квантования, установленным на значение 1 с тем, чтобы передать получаемые таким образом коэффициентные данные на схему НК 17 и схему 18 инверсного квантования.

Схема 17 НК выполняет кодирование с переменной длиной квантованных данных совместно с шагом квантования, параметрами прогнозирования и вектором движения и передает закодированные с переменной длиной данные на средство 30 управления кодированием в качестве первого потока бит.

Средство 30 управления кодированием содержит, как видно из фиг. 1, счетчик 31, служащий для подсчета в каждый задаваемый момент или отрезок времени количества кодированных с переменной длиной данных, поступающих со схемы 17 НК, и вычислительную схему 32, служащую для вычисления (определения) задаваемого (назначаемого) количества кодов на единицу времени на основе количества данных с выхода счетчика 31 и общего количества данных, пригодных к использованию. Счетчик 31 подсчитывает количество данных первого потока бит в каждый заранее задаваемый отрезок времени, например в течение каждого одного кадра, с тем, чтобы вычислять (определять) порог (предел насыщенности) каждого кадра и передать это пороговое значение на схему 32 вычисления скорости следования бит.

Схема 32 вычисляет (определяет) количество кодов, назначаемое для каждого кадра, т.е. среднюю скорость кодирования за время каждого кадра на основе предельной емкости кадра и общего количества пригодных к использованию (используемых) данных, и затем передает значение этого назначаемого количества кодов в схему 33 установки шкалы квантования из состава второго средства кодирования 40.

По существу, схема 32 вычисления скорости кодирования выполняет следующие расчеты. Пусть, например, количество всех кадров есть величина N, общее количество используемых данных есть B, предельная емкость (кодовая емкость) i-го, (i = 0, 1, 2,... N-1) кадра есть d_i и количество назначаемых кодов до i-го кода есть b_i, то когда в виде условия задается, что назначаемая кодовая величина b_i должна быть пропорциональна предельной кодовой емкости d_i, как будет указано в формуле (1), то общее количество данных B может быть вычислено путем сложения задаваемого количества кодов всех кадров, как указано ниже в формуле (2). Величина "a" в следующих формулах есть константа.

$$b_i = a \times d_i \quad ...(1)$$

$$B = \sum_{i=0} bi = \sum_{i=0} aXdi = a \quad o \quad (2)$$

Соответственно, эта константа "a" может быть вычислена с помощью следующей формулы (3). После подстановки этой константы "a" в формулу (1) назначаемое количество кодов b_i до (относительно) i-го кадра может быть вычислено по формуле (4)

$$a = \frac{B}{\sum_{i=0}^{N-1}} \quad (3)$$

$$bi = di \times \frac{B}{\sum_{i=0}^{N-1}} \quad (4)$$

Таким образом, схема 32 вычисления скорости следования битов увеличивает кодовую величину b_i для кадров изображения сложного вида и уменьшает эту величину для кадров изображения более простого вида.

С другой стороны, схема 18 выполняет инверсное квантование уже квантованных в обычном режиме данных, поступающих со схемы 16 с шагом квантования, который устанавливают равным 1 для воспроизведения коэффициентных данных (при этом добавляются искажения, вызванные квантованием), с выхода схемы 15 ДКП и передачи затем этих данных на схему 20 ИДКП.

Схема 20 ИДКП выполняет обработку ИДКП коэффициентных данных для воспроизведения входных данных изображений, соответствующих выходным данным схемы прогнозирующего кодирования 14 во внутрикадровом режиме, и воспроизводит разностные данные во

RU 2123769 C1

RU 2123769 C1

внутрикадровом/вперед/назад/в обоих направлениях прогнозирующем режимах, передавая затем воспроизведенные данные на схему 21 суммирования.

Если установлен режим прогнозирования "вперед/назад/ в обоих направлениях", на схему 21 суммирования поступают прогнозируемые данные изображений с компенсацией движения со схемы 23 компенсации движения. Схема 21 суммирует эти прогнозированные данные изображения и разностные данные, переданные со схемы 20 ИДКП, в результате чего воспроизводятся данные изображения, соответствующие входным данным изображения.

Воспроизведенные таким образом данные изображения хранятся в ЗУ 22 кадровой памяти. По существу, схема 18 инверсного квантования и суммирующая схема 21 образуют локальную схему декодирования, обрабатывающую квантованные данные с выхода схемы 16 квантования на основе заданного режима прогнозирования с тем, чтобы записать полученные декодированные данные в ЗУ 22 кадровой памяти в качестве вперед или назад прогнозированных изображений. Это ЗУ 22 состоит из множества более мелких ячеек кадровой памяти. Здесь производится переключение банка кадровой памяти. В зависимости от параметров данных, подлежащих кодированию, один очный кадр, например, выдается на выход в качестве прогнозированных вперед или назад данных изображения. Далее, в случае прогнозирования "вперед/назад/ в обоих направлениях" прогнозируемые вперед данные изображения и прогнозируемые назад данные изображения, например, усредняются и эти усредненные данные выдаются на выход. Эти прогнозированные данные изображений представляют собой точно такие же данные, как и данные изображений, воспроизведенные устройством декодирования изображения, которое будет описано ниже. На основе этого прогнозированного изображения следующее подлежащее обработке изображение подвергается кодированию в режиме "вперед/назад/в обоих направлениях".

Ниже описывается работа второго кодирующего средства 40. Здесь следует отметить, что поскольку электронные схемы, за исключением схемы 33 установления шкалы квантования, элемента задержки 43, схемы квантования 46 и буфера передачи 49, образующие второе средство кодирования 40, выполняют те же самые операции, что и аналогичные схемы описанного ранее первого средства кодирования 10, то их описание опущено.

Элемент задержки 43 задерживает входные данные изображения, например, до того момента, когда сигнал управления кодированием пройдет на выход средства управления кодированием 30. Затем в схеме прогнозирующего кодирования 44 и схеме 45 ДКП задержанные входные данные изображения подвергаются обработке в виде прогнозируемого кодирования и ДКП, которые выполняются в соответствии в режимом прогнозирования, задаваемым схемой 13, работающей в режимах прогнозирования "внутрикадровом/вперед/назад/в обоих направлениях", и выполняющей функцию принятия решения. Таким образом вырабатываются коэффициентные данные.

Схема 33 установки шкалы квантования определяет (вычисляет) количество кодов, назначаемых для каждого макроблока (например, величину, получаемую путем деления количества кодов, назначаемых для каждого кадра, на число макроблоков в одном кадре) из переданного количества кодов для каждого кадра для выполнения сравнения между количеством кодов, сформированных в макроблоке, который обнаруживается в обратной связи буфера передачи 49, и количеством кодов, назначаемых для каждого макроблока. Работа схемы 33 установки шкалы квантования состоит в том, что она таким образом задает скорости следования битов при кодировании соответствующих кадров, чтобы они были как можно ближе к средним значениям скорости следования битов при кодировании за время каждого кадра. В том случае, когда количество кодов, сформированных в соответствующем макроблоке оказывается больше, чем количество назначенных (заданных) кодов для каждого макроблока, схема 33 устанавливает шаг квантования для следующего макроблока на большую величину для того, чтобы уменьшить количество кодов, которые будут сформированы в следующем макроблоке, и напротив, если количество кодов, сформированных в соответствующем макроблоке, оказывается меньше, чем количество кодов, заданное для каждого макроблока, эта схема 33 устанавливает уменьшенную величину шага квантования для следующего макроблока для увеличения числа формируемых кодов. Здесь следует отметить, еще одну функцию схемы 33, которая состоит в том, что когда в цепи обратной связи буфера передачи 49 появляются указания о том, что буфер 49 находится в состоянии, близком к переполнению, то эта схема 33 также увеличивает шаг квантования для предотвращения переполнения независимо от результатов сравнения заданного и формируемого фактически количества кодов. С другой стороны, если контур обратной связи буфера передачи 49 сигнализирует о том, что содержимое буфера близко к нижнему пределу, схема 33 установления шкалы квантования уменьшает величину шага квантования для увеличения числа формируемых кодов также независимо от результатов сравнения заданного и фактически формируемого количества кодов. Хотя выше было сказано, что сравнение количества заданных (назначаемых) кодов и фактически формируемых выполняется в каждом макроблоке для переключения величины шага квантования для каждого такого макроблока, подобное же переключение может быть выполнено и для каждой секции. Кроме того, хотя выше было указано, что количество фактически формируемых кодов определяют из содержимого буфера передачи 40, эти же показатели можно получить непосредственно с выхода схемы 47 неравномерного кодирования (с переменной длиной). Схема 33 установления шкалы квантования именно таким образом передает величину шага квантования в схему 46 квантования.

Эта схема квантования 46 квантует

RU 2123769 C1

RU 2123769 C1

коэффициентные данные, получаемые от схемы 45 ДКП, с шагом квантования, величина которого передается со схемы 33 установки шкалы для вырабатывания квантованных данных.

Схема НК 47 дает возможность подвергнуть квантованные данные, приходящие со схемы 46 квантования, неравномерному кодированию (с переменной длиной) с шагом квантования, величина которого передается со схемы 33 установки шкалы квантования, и в режиме прогнозирования "внутрикадрового/вперед/назад/в обоих направлениях", задаваемого схемой 13 прогнозируемого решения, а также с учетом вектора движения со схемы его обнаружения 11, после чего полученные таким образом данные кодирования с переменной длиной пересылаются в буфер 49 передачи как второй поток бит.

В настоящем варианте устройства кодирования изображений, как показано на фиг. 3, после ввода данных изображений через элемент задержки 43 в шаге 1 алгоритма, схема 33 установки шкалы получает от средства 30 управления кодированием значение количества кодов, задаваемых для кадра, подлежащего текущему кодированию в шаге 2 алгоритма, после чего обработка переходит к шагу 3.

В шаге 3 алгоритма схема 44 прогнозируемого кодирования совместно со схемой 47 НК выполняют обработку данных изображения в виде прогнозируемого кодирования и ДКП обработки, а затем квантуют коэффициентные данные с шагом квантования, задаваемым на основе количества задаваемых кодов, что дает возможность выполнить кодирование с переменной длиной. Затем обработка переходит к шагу 4.

В шаге 4 алгоритма, независимо от того, завершено или нет кодирование всех кадров (последовательности их), которые имеют один и тот же размер изображения или обрабатываются при одинаковой скорости передачи бит, принимается решение об окончании алгоритма. Если обработка все же на завершена, алгоритм возвращается к шагу 1. Таким образом реализуется кодирование с переменной скоростью, когда скорость кодирования изменяется в рамках массивов кадров. Соответственно, даже если изображения (кадры) сложного вида поступают в виде последовательности, невозможно обеспечить столь же большую величину шага квантования, как в обычных (известных) устройствах. Так обеспечивается однородное и высокое качество изображения во всех возможных ситуациях.

Буфер передачи 49 временно хранит данные кодирования с переменной длиной, чтобы затем выдать их на считывание с фиксированной скоростью следования бит, обеспечивая упорядоченный вывод кодированных с переменной длиной данных в виде потока бит. Поток бит с выхода буфера передачи 49 мультиплексируется, например, с кодированным аудиосигналом, сигналом синхронизации и т. п. Кроме того, сюда же добавляется код коррекции ошибок и модуляция с заранее заданными параметрами, пригодными для передачи или регистрации данных. Затем обработанный таким образом поток бит передается на устройство декодирования изображения либо через среду распространения, либо регистрируется на носителе записи изображения 55, представляющем собой оптический диск, магнитный диск или магнитную ленту, что уже упоминалось выше и показано на фиг. 1. Поскольку во втором средстве кодирования 40 производится такая обработка в виде кодирования с переменной скоростью, когда количество задаваемых кодов увеличивается при обработке изображений сложного вида и уменьшается при обработке изображений простого вида, то нет необходимости поддерживать высокую и фиксированную скорость во всех случаях во избежание значительного ухудшения качества изображений сложного вида, как это делается в обычных (известных) устройствах. Благодаря этому время регистрации изображения на носителе 55 может быть продлено.

С другой стороны, схема 48 выполняет инверсное квантование данных, квантованных обычным методом и переданных со схемы 46 квантования, причем величина шага квантования при этом та же, что и в схеме 46, что необходимо для воспроизведения коэффициентных данных (при этом добавляются искажения из-за квантования), соответствующих выходным данным схемы 45 ДКП, после чего эти коэффициентные данные пересылаются в схему 50 ИДКП. Кроме того, схема 48 инверсного квантования и суммирующая схема 51 образуют локальную схему декодирования, декодирующую квантованные данные, поступающие с выхода схемы квантования 46, после чего полученное таким образом декодированное изображение записывается в ЗУ 52 кадровой памяти как прогнозированное вперед или назад изображение. Данные изображения, хранящиеся в ЗУ 52, используются в качестве прогнозированного изображения для следующего подлежащего обработке изображения.

Следует отметить, что хотя в вышеупомянутом варианте реализации изобретения задаваемое количество кодов для заранее задаваемого отрезка времени, т. е. средняя скорость кодирования на протяжении заранее заданного отрезка времени, определяется для каждого кадра, причем за заданный отрезок времени принимается длительность кадра, оно (изобретение) не ограничивается только таким вариантом реализации. Например, в качестве заранее задаваемого отрезка времени может быть использована длительность GOP (Группы Изображений) в нормативах так называемой Экспертной Группы по Подвижным Изображениям (ЭГПИ). Название этой группы ЭГПИ (MPEG) используется и как наименование системы кодирования подвижных изображений, исследованной в Рабочей Группе (РГ) 11 Подкомитета 29 (ПК) Объединенного Технического Комитета (ОТК) в составе Международной Организации по Стандартизации (МОС) и Международного Электротехнического Комитета (МЭТК).

Согласно нормативам (ЭГПИ) группа изобаржений GOP содержит по меньшей мере одно так называемое I-изображение и множество P-изображений или B-изображений

RU 2123769 C1

RU 2123769 C1

(не I-изображения). Рассмотрим практический пример, когда GOP состоит из одиночного I-изображения, четырех P-изображений с периодом следования в 3 изображения и десять B-изображений, а схема управления кодированием 30 определяет количество задаваемых кодов для каждой GOP. Здесь I-изображение подвергается внутриполевому или внутрикадровому кодированию. P-изображение есть изображение, прогнозируемое только в направлении вперед и подлежащее внутриполевому или внутрикадровому кодированию.
B-изображение есть изображение, которое может прогнозироваться в направлениях вперед, назад и в обоих направлениях, оно подлежит внутриполевому или внутрикадровому кодированию.

Если далее предполагается, например, как показано на фиг. 5, что два последовательных произвольных изображения в составе ГИ, состоящем из определенного количества изображений, есть повторяющиеся периодически I-изображение и P-изображение, и величина шага квантования установлена равной, например, 1, то первое средство кодирования 10 может выполнить прогнозирующее кодирование, ДКП обработку и кодирование с переменной длиной данных этих I-изображения и P-изображения для выработки кодированных с переменной длиной данных, которые затем будут переданы на средство 30 управления кодированием. Два эти изображения именно типа I и P выбраны с целью проверки сложности рисунка изображения и степени корреляции между кадрами. По количеству формируемых кодов I-изображения возможно определить (оценить) сложность изображения. По количеству формируемых кодов P-изображения можно определить степень корреляции между кадрами. Поскольку последовательность множества кадров имеет в общем-то аналогичные пиктографические элементы, то даже по двум выбранным изображениям можно оценить тенденцию структуры GOP.

Средство 30 управления кодированием подсчитывает (вычисляет) количество данных I-изображений в каждой ГИ, обозначаемое как I_j бит, и количество данных P-изображений P_j бит, и определяет для каждой GOP фактический предел насыщенности (порог) (количество формируемых кодов d_j, где j = 0, 1, 2,...) на основе величин I_j и P_j, а также количество N P-изображений, составляющих GOP, как указано ниже в формуле (5).

$$\text{GOP } d_j = \text{бит } J_i + N^x \text{ бит } P_j \ldots (5)$$

Средство 30 управления кодированием определяет количество кодов, отведенное для каждой, на основе величины порога каждой GOP и общего количества используемых данных, а затем пересылает это значение количества задаваемых кодов к второму средству кодирования 40.

Если принять, что число групп изображений GOP есть M, общее количество пригодных к использованию данных есть B, задаваемое количество кодов применительно к j-му GOP есть GOP b_j, а также задать, что эта величина GOP b_j пропорциональна пороговому значению, как следует из приведенной ниже формулы (6), то общее количество данных B определяется путем суммирования величин b_i все, как записано в формуле (7). Величина a в формуле (6) есть константа.

$$\text{GOP } b_j = a \times \text{GOP } d_j \ldots (6)$$

$$B = \sum_{j=0}^{M-1} \text{GOP } b_i = \sum_{j=0}^{M-1} a \times \text{GOP } d_j = a \times \sum_{j=0}^{M-1} \text{GOP } d_j \quad (7)$$

Эта константа a может быть определена из формулы (8). Подставляя эту константу в формулу (6), можно по формуле (9) определить задаваемое количество кодов ГИ b_j для j-го ГИ.

$$a = \frac{B}{\sum_{j=0}^{M-1} \text{GOP } d_j} \quad (8)$$

$$\text{GOP } b_j = \text{GOP } d_j \times \left(\frac{B}{\sum_{j=0}^{M-1} \text{GOP } d_j} \right) \quad (9)$$

Итак, средство 30 управления кодированием увеличивает величину количества задаваемых кодов GOP b_j в случае, например, группы изображений, в которых содержатся изображения со сложным рисунком или которые обладают низкой степенью корреляции между кадрами, и напротив, в группе изображений, включающих изображения с простым рисунком или имеющих высокую степень корреляции между кадрами, это средство 30 уменьшает предназначаемое для нее количество кодов.

Когда на второе средство кодирования 40 в шаге 1 алгоритма, приведенного на фиг. 6, через элемент задержки 43 поступают данные изображений, это средство 40 определяет, являются или нет эти данные начальными данными (шаг 2) группы изображений. Если это так, алгоритм обработки переходит к шагу 3. В противном случае алгоритм переходит к шагу 4.

В шаге 3 второе средство кодирования 40 получает от средства управления 30 значение числа кодов, предназначаемых для GOP, подвергаемой текущему кодированию. Затем алгоритм обработки переходит к шагу 4.

В шаге 4 второе кодирующее средство 40 выполняет прогнозирующее кодирование и ДКП-обработку, а также квантует коэффициентные данные с шагом, заданным на основе назначенного числа кодов, что дает возможность выполнить кодирование с переменной длиной. Затем алгоритм обработки переходит к шагу 5.

Здесь схема 33 установки шкалы квантования задает количество кодов для каждого кадра из числа переданного ему общего количества кодов, назначенных для каждой ГИ, принимая во внимание тип изображения (I-изображение, P-изображение, B-изображение) при конкретном кодировании, например тип изображения, показанный на фиг. 4. В практическом аспекте можно сказать, что количество кодов, назначаемых для I-изображений, увеличивается, количество кодов для B-изображений уменьшается, а количество кодов для P-изображений имеет промежуточное между ними значение.

В последующем шаге 5 независимо от того, завершено или нет кодирование всех тех кадров (их последовательности), для которых был задан одинаковый размер изображения

RU 2123769 C1

RU 2123769 C1

или одинаковая скорость передачи бит, принимается соответствующее решение. Если последовательность закончилась, то и обработка считается завершенной. В противном случае алгоритм возвращается к шагу 1. Так реализуется неравномерное кодирование (с переменной длиной), когда скорость кодирования изменяется в каждой GOP. Если даже имеет место последовательность изображений (кадров) со сложным рисунком, то все же нет возможности сделать шаг квантования достаточно большим для таких изображений, как в обычных (известных) устройствах. В результате появляется возможность формирования однородного изображения высокого качества в целом для всех ситуаций. Кроме того, поскольку задаваемое количество кодов для каждой GOP определяется в настоящем варианте устройства по двум изображениям, то по сравнению с вышеописанным вариантом может быть установлена более высокая скорость обработки. Осталось только отметить тот очевидный факт, что количество кодов соответственных групп изображений определяются на основе количества данных всех изображений в составе GOP.

Необходимо отметить, что настоящее изобретение не ограничивается описанными выше вариантами его. Хотя, например, в описанных выше вариантах для преобразовательного кодирования используется метод ДКП, преобразование Страто, преобразование Хаара или преобразование Вавелета также могут быть применены в описанных случаях.

Промышленное применение.

Как должно быть понятно из приведенного выше описания, в настоящем изобретении используется электронная схема и последовательность операций, позволяющие кодировать входной видеосигнал путем выполнения, например, прогнозирующего кодирования, ДКП-обработки и квантования данных с фиксированным шагом квантования, кодирования (с переменной длиной) для выработки первых кодированных данных и последующего определения (вычисления) количества кодов, отводимых (задаваемых) для каждого кадра или каждой GOP, на основе учета количества данных в каждый заранее заданный отрезок времени, например за время длительности каждого кадра или каждой ГИ из первых кодированных данных, и общего количества данных, пригодных для кодирования входного видеосигнала в каждый заранее задаваемый момент на основе назначаемого количества кодов для вырабатывания вторых кодированных данных. Таким образом осуществляется кодирование с переменной скоростью, когда скорость кодирования изменяется в заранее задаваемый отрезок времени. В результате, если даже имеет место последовательность изображений (кадров) сложного вида, нет возможности обеспечить такое большое значение шага квантования для этих изображений, как в обычных устройствах. Благодаря этому можно достигнуть однородного изображения высокого качества в целом в различных случаях.

Кроме того, поскольку вторые кодированные данные, формирование которых описано выше, имеют переменную скорость, то в случае, когда такие данные регистрируются на носителе записанного изображения, ограниченный объем памяти может быть использован более эффективно, а время регистрации изображения на носителе может быть продлено. Наконец, данные изображения, имеющего высокое качество и однородность во всех возможных случаях, могут быть воспроизведены с носителя записанного изображения.

Формула изобретения:

1. Способ кодирования изображений, включающий операции кодирования по меньшей мере части входного видеосигнала для выработки первых кодированных данных, определения скорости кодирования в каждый заранее задаваемый момент или отрезок времени на основе количества первых кодированных данных и общего количества пригодных к использованию данных, отличающийся тем, что предусматривает кодирование входного видеосигнала в каждый заранее задаваемый отрезок времени на основе скорости кодирования для выработки вторых кодированных данных.

2. Способ по п.1, отличающийся тем, что по меньшей мере часть входного видеосигнала квантуют с фиксированным шагом квантования для выработки первых кодированных данных.

3. Способ по п.1, отличающийся тем, что общее количество пригодных к использованию данных пропорционально распределяют в соответствии с требуемым количеством данных в каждый заранее задаваемый отрезок времени для определения таким образом скорости кодирования в каждый заранее задаваемый отрезок времени.

4. Способ кодирования изображений, включающий операции заранее задаваемого прогнозирующего кодирования или заранее задаваемого преобразовательного кодирования по меньшей мере части входного видеосигнала для вырабатывания первых коэффициентных данных, квантования первых коэффициентных данных с фиксированным шагом квантования для вырабатывания первых квантованных данных, обеспечивая возможности обработки первых квантованных данных в виде неравномерного кодирования (с переменной длиной) для формирования первого потока бит, определения скорости кодирования в каждый заранее задаваемый отрезок времени на основе количества данных в первом потоке бит и общего количества пригодных к использованию данных, отличающийся тем, что прогнозирующее кодирование и/или заранее задаваемое преобразовательное кодирование входного видеосигнала для генерации вторых коэффициентных данных, квантование вторых коэффициентных данных с шагом квантования, задаваемым на основе скорости кодирования в каждый заранее задаваемый отрезок времени для выработки вторых квантованных данных, и обеспечение возможности кодирования с переменной длиной вторых квантованных данных для формирования второго потока бит.

5. Способ по п.4, отличающийся тем, что скорость кодирования определяют в каждом кадре на основе количества данных в каждом кадре первого потока бит и общего количества пригодных к использованию

RU 2123769 C1

RU 2123769 C1

данных.

6. Способ по п.4, отличающийся тем, что скорость кодирования определяют в каждой группе изображений (ГИ) на основе по крайней мере количества части данных каждой группы изображений (ГИ), состоящей из множества кадров из первого потока бит, и на основе общего количества пригодных к использованию данных.

7. Способ по п.6, отличающийся тем, что скорость кодирования в каждой группе изображений (ГИ) определяют на основе количества данных в изображении с внутрикадровым кодированием и изображения, кодированного с прогнозированием вперед в составе группы изображений (ГИ).

8. Способ по п. 4, отличающийся тем, что общее количество пригодных к использованию данных распределяют пропорционально в зависимости от количества данных в первом потоке бит в каждый заранее задаваемый отрезок времени для определения скорости кодирования в каждый заранее задаваемый отрезок времени.

9. Способ кодирования изображений, включающий операции определения предела кодирования в каждом блоке изображения во входном видеосигнале, установки скорости кодирования в каждом заранее задаваемом блоке изображения на основе предела кодирования и общего количества данных, пригодных к использованию, отличающийся тем, что предусматривает кодирование входного видеосигнала таким образом, что скорости кодирования соответственных блоков изображения поддерживают в соответствии с установленной для каждого блока изображения скоростью кодирования.

10. Способ по п.9, отличающийся тем, что в качестве заранее задаваемого блока изображения задают кадр.

11. Способ по п.9, отличающийся тем, что в качестве заранее задаваемого блока изображения задают группу изображений (ГИ), состоящую из множества кадров.

12. Способ по п.9, отличающийся тем, что заранее задаваемому прогнозирующему кодированию и/или заранее задаваемому преобразовательному кодированию подвергают по крайней мере часть входного видеосигнала для выработки коэффициентных данных с тем, чтобы затем произвести квантование этих коэффициентных данных с фиксированным шагом квантования и определить таким образом предел кодирования.

13. Устройство кодирования изображений, содержащее первое средство кодирования по крайней мере части входного видеосигнала для вырабатывания первых кодированных данных, средство управления кодированием, служащее для определения скорости кодирования в заранее задаваемый отрезок времени на основе количества первых кодированных данных в каждый заранее задаваемый отрезок времени, поступающих от первого средства кодирования, и общего количества данных, пригодных для использования, отличающееся тем, что содержит второе средство кодирования входного видеосигнала в каждый заранее задаваемый отрезок времени на основе скорости кодирования в каждый заранее задаваемый отрезок времени, задаваемой средством управления кодированием, для выработки вторых кодированных данных.

14. Устройство по п.13, отличающееся тем, что первое средство кодирования содержит средство для квантования по крайней мере части входного видеосигнала с фиксированным шагом квантования.

15. Устройство по п.13, отличающееся тем, что средство управления кодированием выполняют функцию пропорционального распределения общего количества данных в зависимости от наличия данных в каждый заранее задаваемый отрезок времени, определяя таким образом скорость кодирования в каждый заранее задаваемый отрезок времени.

16. Устройство кодирования изображений, содержащее первое средство для выполнения заранее задаваемого прогнозирующего кодирования и/иди заранее задаваемого преобразовательного кодирования по меньшей мере части входного видеосигнала для выработки первых коэффициентных данных, первое средство квантования первых коэффициентных данных с фиксированным шагом квантования для образования первых квантованных данных, первое средство неравномерного кодирования, обеспечивающее такую обработку квантованных данных, приходящих с первого средства квантования, для формирования первого потока бит, средство управления кодированием для определения скорости кодирования в каждый заранее задаваемый отрезок времени на основе количества данных в первом потоке бит, поступающем с первого средства неравномерного кодирования, и общего количества используемых данных, отличающееся тем, что содержит второе средство кодирования для выполнения заранее задаваемого прогнозирующего кодирования и/или заранее задаваемого прогнозирующего преобразовательного кодирования входного видеосигнала для выработки вторых коэффициентных данных, второе средство квантования для квантования вторых коэффициентных данных, поступающих со второго кодирующего средства, с шагом квантования, задаваемым с учетом скорости кодирования средством управления кодированием в каждый заранее задаваемый отрезок времени для вырабатывания вторых квантованных данных, и второе средство неравномерного кодирования, обеспечивающее такую обработку вторых квантованных данных, поступающих со второго средства квантования для формирования второго потока бит.

17. Устройство по п.16, отличающееся тем, что средство управления кодированием определяет скорость кодирования в каждом кадре на основе количества данных в каждом кадре первого потока бит и общего количества пригодных данных.

18. Устройство по п.16, отличающееся тем, что средство управления кодированием определяет скорость кодирования в каждой группе изображений (ГИ) на основе количества данных по меньшей мере в части каждой группы изображений (ГИ), состоящей из множества кадров первого потока бит, а также общего количества пригодных к

RU 2123769 C1

RU 2123769 C1

использованию данных.

19. Устройство по п.18, отличающееся тем, что средство управления кодированием определяет скорость кодирования в каждой группе изображений (ГИ) на основе количества данных в изображениях внутрикадрового кодирования и в изображениях, кодированных с прогнозированием вперед в составе группы изображений (ГИ).

20. Устройство по п.16, отличающееся тем, что средство управления кодированием пропорционально распределяет общее количество пригодных к использованию данных в зависимости от количества данных в первом потоке бит в каждый заранее задаваемый отрезок времени, определяя таким образом скорость кодирования в каждый заранее задаваемый отрезок времени.

21. Устройство кодирования изображений, содержащее средство вычисления (определения) предела кодирования в каждом заранее задаваемом блоке изображений во входном видеосигнале, средство установки скорости кодирования в каждом заранее задаваемом блоке изображений на основе предельных параметров кодирования, получаемых от средства вычисления этих предельных параметров, с учетом общего количества пригодных к использованию данных, отличающееся тем, что содержит средство кодирования, обеспечивающее возможность такого кодирования входного видеосигнала, когда скорости кодирования соответственных блоков изображения соответствуют скорости кодирования каждого блока изображений, установленной средством установки скорости кодирования.

22. Устройство по п.21, отличающееся тем, что средство вычисления предела кодирования определяет эти параметры кодирования для каждого кадра.

23. Устройство по п.21, отличающееся тем, что средство вычисления предельных параметров кодирования определяет эти параметры для каждой группы изображений (ГИ), состоящей из множества кадров.

24. Устройство по п.21, отличающееся тем, что средство вычисления предела кодирования выполняет заранее задаваемое прогнозирующее кодирование и/или заранее задаваемое преобразовательное кодирование по крайней мере части входного видеосигнала для выработки коэффициентных данных и их последующего квантования с фиксированным шагом, в результате чего вычисляют (определяют) предел кодирования.

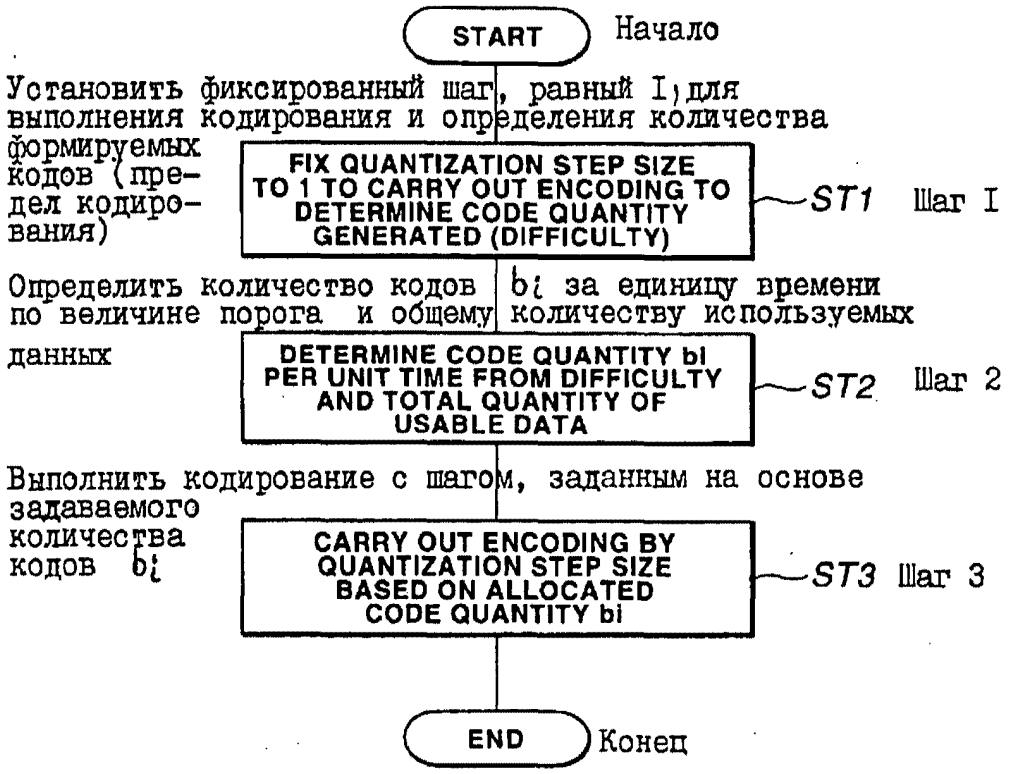
25. Носитель информации для записи изображений, на котором записан второй поток бит, полученный путем кодирования по крайней мере части входного видеосигнала для выработки первых кодированных данных, определения скорости кодирования в каждый заранее задаваемый отрезок времени на основе количества первых кодированных данных в каждый заранее задаваемый отрезок времени, а также общего количества используемых данных, отличающийся тем, что второй поток бит получен в результате кодирования входного видеосигнала в каждый заранее задаваемый отрезок времени на основе скорости кодирования.

26. Носитель информации для записи изображений, на котором записан второй поток бит, полученный путем выполнения заранее задаваемого прогнозирующего кодирования и/или заранее задаваемого преобразовательного кодирования по меньшей мере части входного видеосигнала, выработки первых коэффициентных данных и их квантования с фиксированных шагом для выработки первых квантованных данных, обеспечивания возможности неравномерного кодирования первых квантованных данных для формирования первого потока бит, определения скорости кодирования в каждый заранее задаваемый отрезок времени на основе количества данных в первом потоке данных и общего количества пригодных к использованию данных, отличающийся тем, что второй поток бит получен в результате выполнения заранее задаваемого прогнозирующего кодирования и/или заранее задаваемого преобразовательного кодирования входного видеосигнала для выработки вторых коэффициентных данных, квантования вторых коэффициентных данных с шагом квантования, задаваемым на основе скорости кодирования в каждый заранее задаваемый отрезок времени для выработки вторых квантованных данных и обеспечения возможности неравномерного кодирования этих вторых квантованных данных.

27. Носитель информации для записи изображений, на котором записаны кодированные данные, полученные путем определения предела кодирования в каждом заранее задаваемом блоке изображения входного видеосигнала, установления скорости кодирования в каждом заранее задаваемом блоке изображения, на основе величины предела кодирования и общего количества пригодных к использованию данных, отличающийся тем, что записанные данные получены в результате кодирования входного видеосигнала таким образом, что скорости кодирования в каждом из соответственных блоков изображения соответствуют скорости кодирования, установленной для каждого блока изображения.

RU 2123769 C1

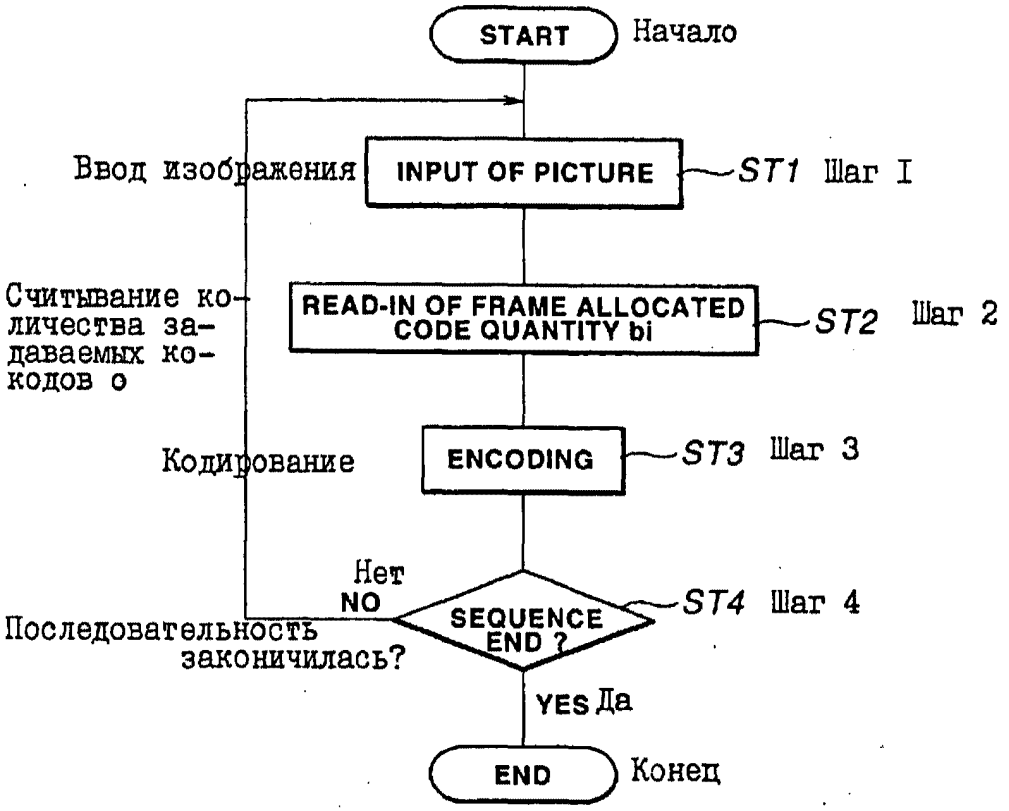
RU 2123769 C1



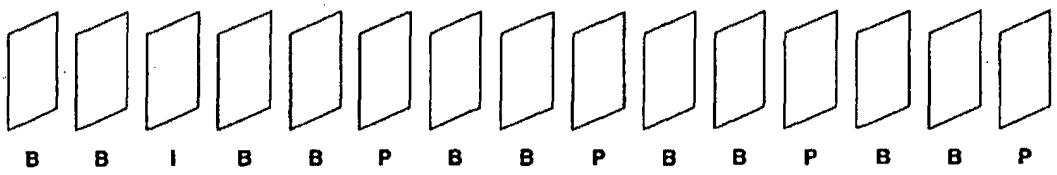
Фиг.2

RU 2123769 C1

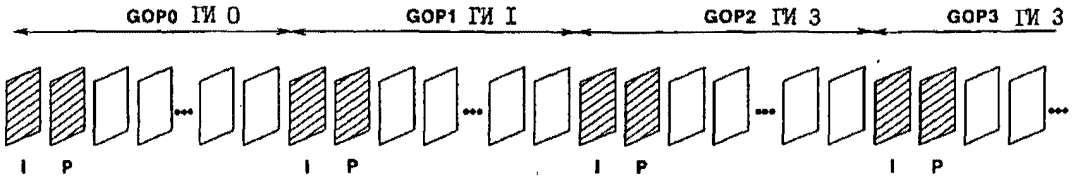
RU 2123769 C1



Фиг.3



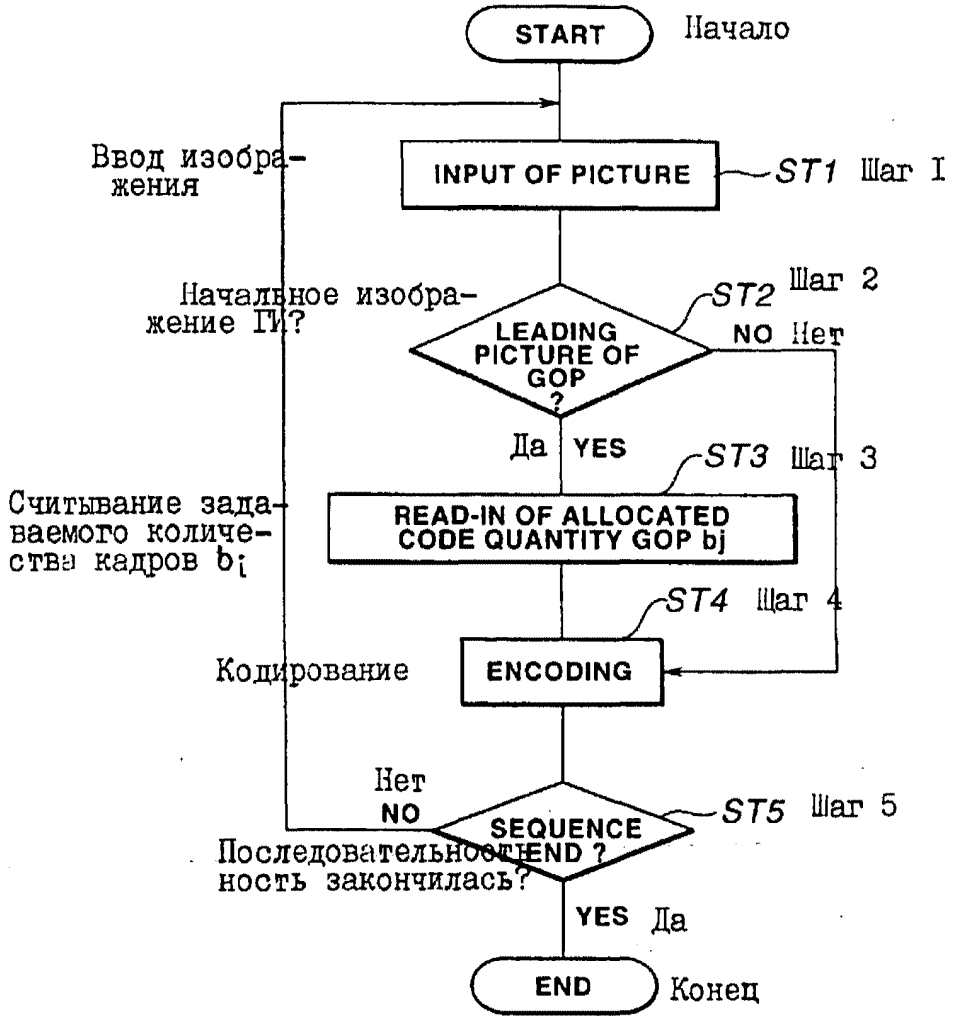
Фиг.4



Фиг.5

RU 2123769 C1

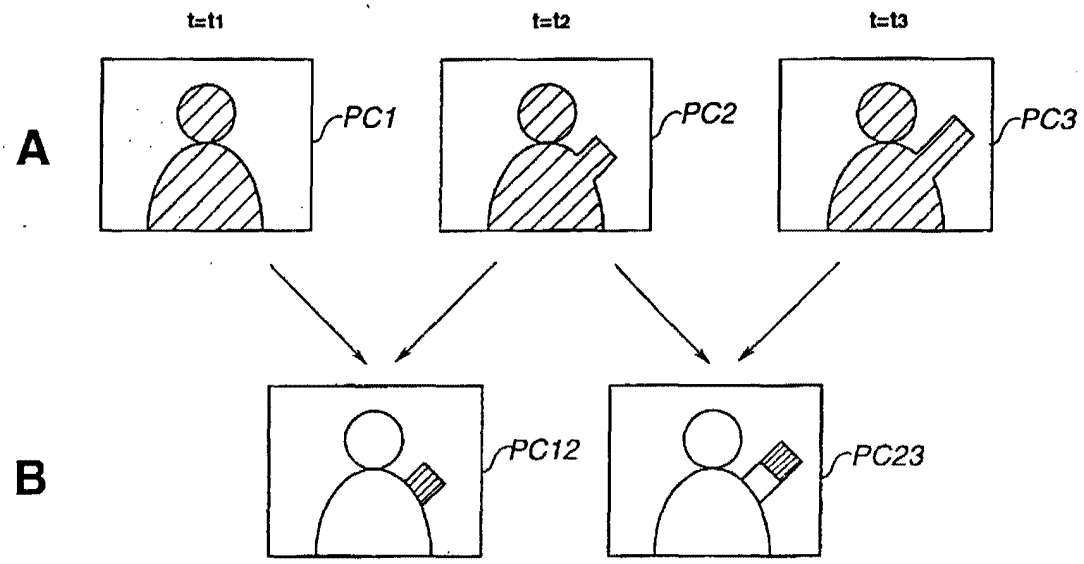
RU 2123769 C1



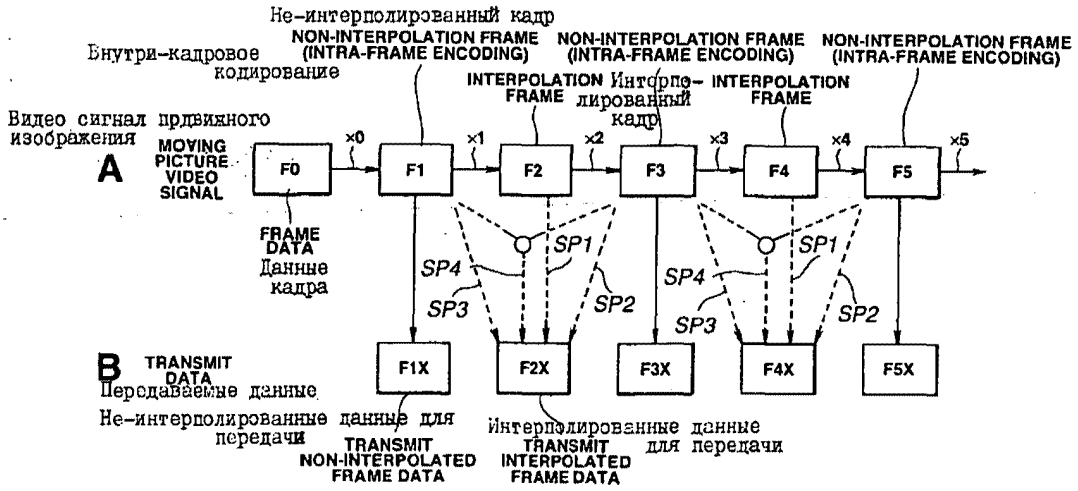
Фиг.6

RU 2123769 C1

RU 2123769 C1



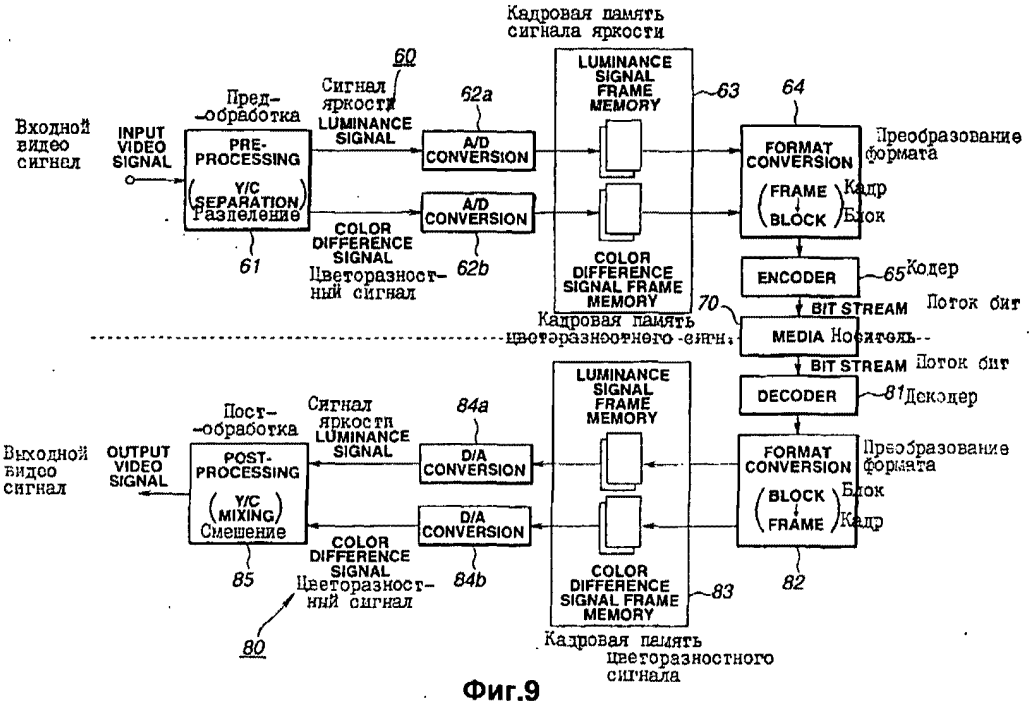
Фиг.7



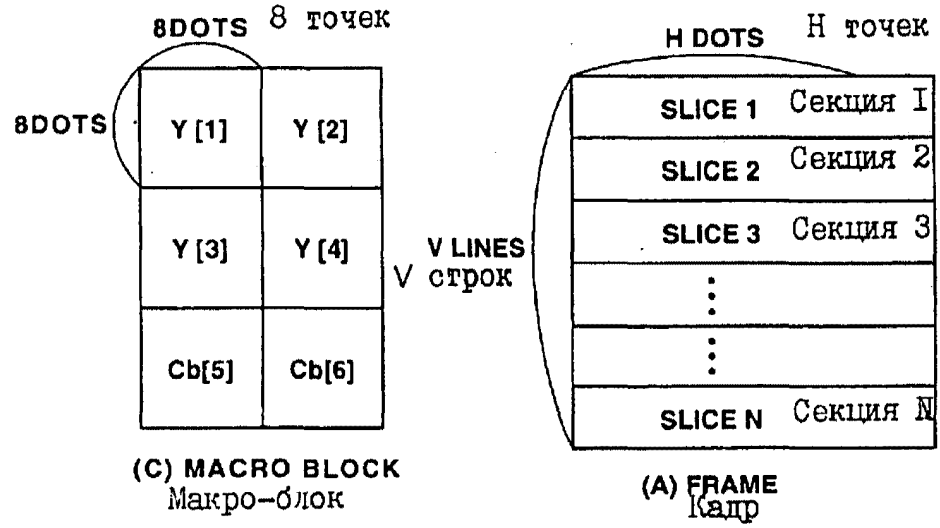
Фиг.8

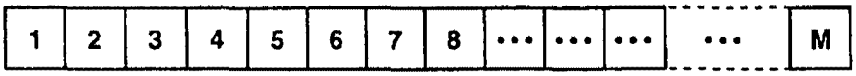
RU 2123769 C1

RU 2123769 C1



Фиг.9



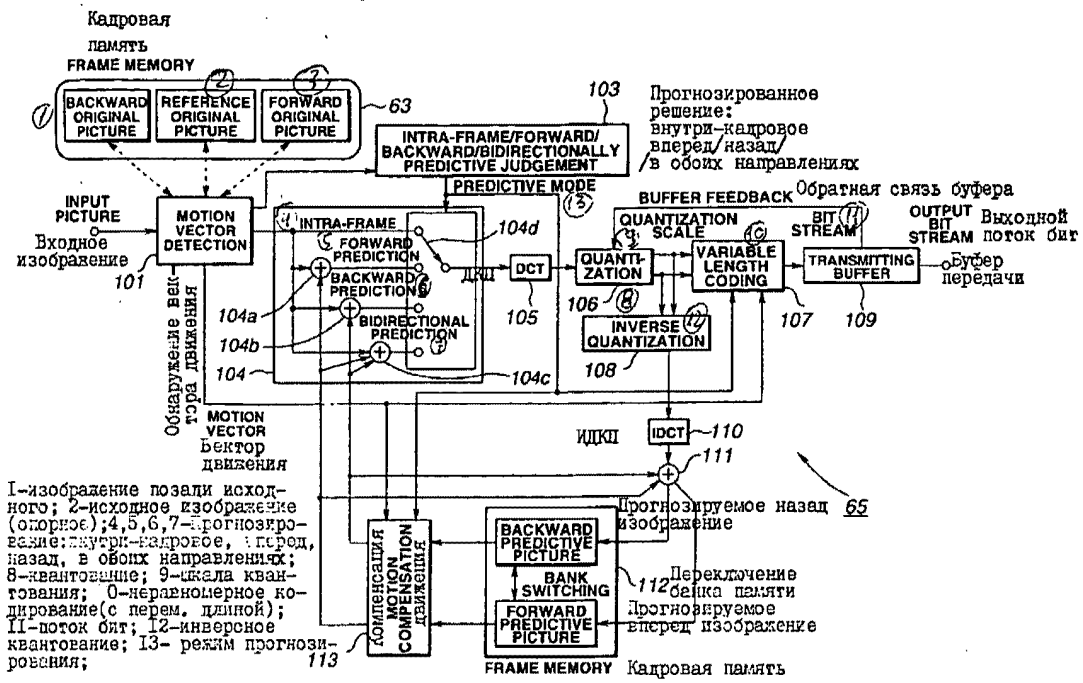


(B) SLICE Секция B

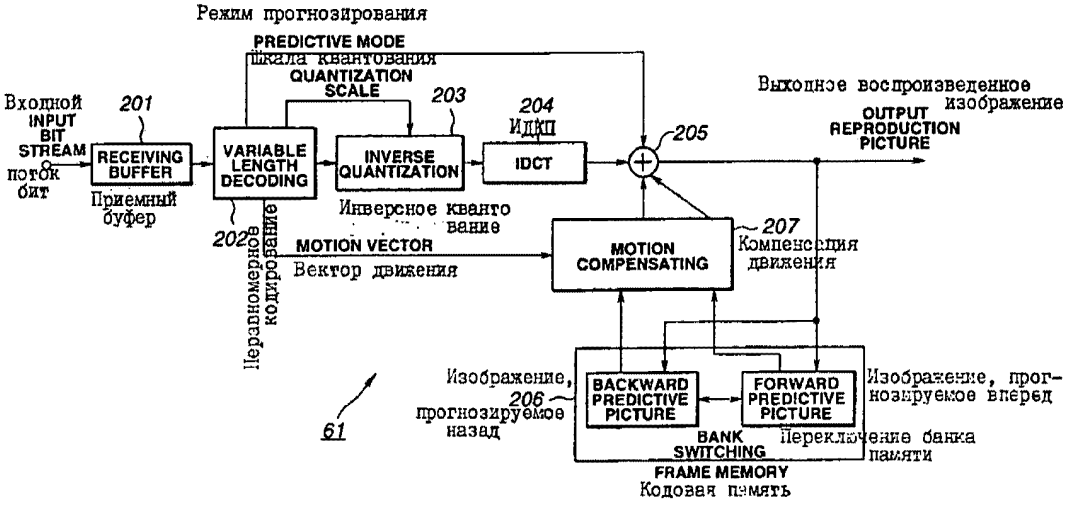
Фиг.10

RU 2123769 C1

RU 2123769 C1



Фиг.11



Фиг.12

RU 2123769 C1

RU 2123769 C1

GHE 1600
OCT 22 2007
PATENT & TRADEMARK OFFICE

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant:	HANNUKSELA	Examiner:	Unassigned
Serial No.:	11/338,934	Group Art Unit:	2613
Filed:	January 25, 2006	Docket No.:	TPO.004.C2
Title:	GROUPING OF IMAGE FRAMES IN VIDEO CODING		

CERTIFICATE UNDER 37 CFR 1.8: The undersigned hereby certifies that this correspondence and the papers, as described hereinabove, are being deposited in the United States Postal Service, as first class mail, in an envelope addressed to: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450 on October 17, 2007.

By: Rennae Johnson

SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT (37 C.F.R. § 1.97(b))

Mail Stop Amendment
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

With regard to the above-identified application, the items of information listed on the enclosed Form 1449 are brought to the attention of the Examiner.

This statement should be considered because it is submitted before the mailing date of a first Office Action on the merits for the above-identified application. Accordingly, no fee is due for consideration of the items listed on the enclosed Form 1449.

These items were also first cited in a communication from a foreign patent office regarding a counterpart foreign application not more than three months prior to the filing of this information disclosure statement.

In accordance with 37 C.F.R. § 1.98(a)(2), only a copy of each foreign document or non-U.S. patent/application listed on the enclosed Form 1449 is provided.

Please note that any notations or markings on the attached document do not reflect particular relevance, or lack thereof, to the present application, nor were they necessarily made by anyone affiliated with the prosecution of the present application.

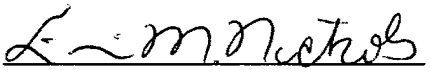
No representation is made that a reference is "prior art" within the meaning of 35 U.S.C. §§ 102 and 103 and Applicant reserves the right, pursuant to 37 C.F.R. § 1.131 or otherwise, to establish that the reference(s) are not "prior art." Moreover, Applicant does not represent that a reference has been thoroughly reviewed or that any relevance of any portion of a reference is intended.

Consideration of the items listed is respectfully requested. Pursuant to the provisions of MPEP § 609, it is requested that the Examiner return a copy of the attached Form 1449, marked as being considered and initialed by the Examiner, to the undersigned with the next official communication.

Respectfully submitted,

HOLLINGSWORTH & FUNK, LLC
8009 34th Avenue South, Suite 125
Minneapolis, MN 55425
952.854.2700

Date: October 17, 2007

By: 
Erin M. Nichols
Reg. No. 57,125

FORM 1449* SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT IN AN APPLICATION	Docket Number: TAPA.004.C2	Application Number: 11/338,934
	Applicant: HANNUKSELA	
	Filing Date: Jan. 25, 2006	Group Art Unit: 2613

OIPE IAP17 OCT 22 2007 PATENT & TRADEMARK OFFICE

U.S. PATENT DOCUMENTS

EXAMINER INITIAL	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	5,144,426	09.01.1992	Tanaka et al.			

FOREIGN PATENT DOCUMENTS

	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	TRANSLATION	
						YES	NO
	01/84850	8 Nov 2001	WO			NA	

OTHER DOCUMENTS (Including Author, Title, Date, Pertinent Pages, Etc.)

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609; draw line through citation if not in conformance and not considered. Include copy of this form for next communication to the Applicant.

*Substitute Disclosure Statement Form (PTO-1449) Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
8 November 2001 (08.11.2001)

PCT

(10) International Publication Number
WO 01/84850 A1

(51) International Patent Classification[7]: H04N 7/50

(21) International Application Number: PCT/EP01/04319

(22) International Filing Date: 13 April 2001 (13.04.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
00201505.5 27 April 2000 (27.04.2000) EP
00203828.9 2 November 2000 (02.11.2000) EP

(71) Applicant: KONINKLIJKE PHILIPS ELECTRONICS N.V. [NL/NL]; Groenewoudseweg 1, NL-5621 BA Eindhoven (NL).

(72) Inventors: BRULS, Wilhelmus, H., A.; Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL). KLEIN GUNNEWIEK, Reinier, B., M.; Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL).

(74) Agent: STEENBEEK, Leonardus, J.; Internationaal Octrooibureau B.V., Prof Holstlaan 6, NL-5656 AA Eindhoven (NL).

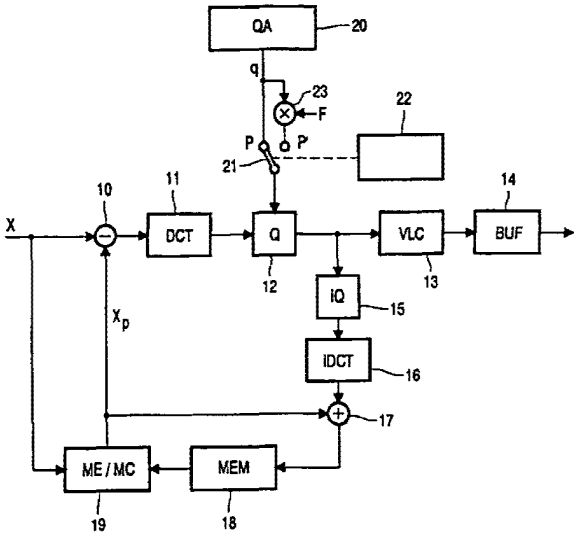
(81) Designated States *(national)*: CN, JP, KR.

(84) Designated States *(regional)*: European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR).

Published:
— *with international search report*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: VIDEO COMPRESSION



(57) Abstract: The concept of B-frames gives the MPEG video compression standard its high encoding efficiency. However, B-frame encoding roughly doubles the complexity of an MPEG encoder. In view thereof, MPEG encoders have been developed which produce I-frames and P-frames only. They are less complex but also less efficient. To improve the efficiency of such "IPP encoders", selected P-frames are quantized more coarsely than other P-frames, for example, by multiplying the conventional quantization step size by 1.4. Although this results in isolated frames ("virtual B-frames") being encoded with a lower quality, the overall perceptual quality is not affected. It has been found that the gain in bit rate obtained by the coarser quantization is not lost in subsequent P-frames, even though the subsequent frames are encoded with reference to the lower quality frames.

WO 01/84850 A1

Video compression

FIELD OF THE INVENTION

The invention relates to a method of compressing a video signal, the method comprising predictively encoding frames of said video signal with reference to a prediction frame, calculating a quantization parameter for each encoded frame, and quantizing the encoded frames in accordance with said quantization parameter. The invention also relates to a compression arrangement, to a transmission or recording method and arrangement, to the compressed video signal and to a storage medium comprising that signal.

BACKGROUND OF THE INVENTION

A video compression method as defined in the opening paragraph has been standardized by the Motion Frames Expert Group and is well-known as MPEG1 or MPEG2. The known method includes transformation of video pixels into frequency coefficients, quantization of said coefficients, and variable-length coding of the quantized coefficients. The quantization is controlled so as to achieve a desired quality or bit rate of the compressed signal.

The MPEG compression method produces I, P and B-frames. I-frames are encoded autonomously, i.e. without reference to another frame. P-frames are predictively encoded with reference to a previous (possibly motion-compensated) I or P-frame. B-frames are bidirectionally predictively encoded with reference to a previous and a subsequent I or P frame. B-frames are not themselves used as reference for encoding other frames.

The concept of B-frames in MPEG provides maximum encoding efficiency. However, the use of B-frames roughly doubles the complexity, memory capacity and memory bandwidth. In view thereof, MPEG encoders have been developed which produce I and P-frames only ("IP encoders"). A disadvantage of IP encoders is their efficiency. They need approximately 10-20% more bit rate than IPB encoders.

OBJECT AND SUMMARY OF THE INVENTION

It is an object of the invention to provide an arrangement and method which overcomes the above-mentioned disadvantage of prior-art IP encoders.

To this end, the invention provides a video compression arrangement and method, a compressed signal, a storage medium, and a transmission or recording method and arrangement as defined in the independent claims. Advantageous embodiments are defined in the dependent claims.

The method in accordance with the invention quantizes selected P-frames more coarsely than other P-frames. This reduces the bit cost but degrades the image quality of said frames. The invention has a surprising effect. It was expected that the corresponding gain in bit cost would be lost in subsequent P-frames because the lower quality frames are used as prediction for subsequent P-frames. However, experiments have shown that this is not the case. It has been found that an IPPPP.. sequence of frames, in which the quantization parameter of every other P-frame is multiplied by a factor of 1.4, has substantially the same bit rate as a conventional IBPBP.. sequence having the same perceptual visual quality. In view thereof, the lower quality P-frames are also referred to as "virtual B-frames".

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 shows a schematic diagram of a preferred embodiment of an arrangement for compressing a video signal encoder in accordance with the invention;

Figs. 2A and 2B show diagrams illustrating the performance of the arrangement in accordance with the invention compared with the performance of a prior-art arrangement;

Fig. 3 shows a block diagram of embodiments of arrangements for transmitting and receiving a video signal; and

Fig. 4 shows a block diagram of embodiments of arrangements for recording a video signal on a storage medium and for playing back from the storage medium.

DESCRIPTION OF EMBODIMENTS

Fig. 1 shows a schematic diagram of an MPEG encoder in accordance with the invention. The Figure shows the encoder in the state in which P-frames are encoded. The encoder is a conventional MPEG encoder in the sense that it comprises a subtraction circuit **10**, a discrete cosine transformer (DCT) **11**, a quantizer (Q) **12**, a variable-length coder (VLC) **13**, a buffer (BUF) **14**, an inverse quantizer (iQ) **15**, an inverse discrete cosine transformer (iDCT) **16**, an adder **17**, a frame memory (MEM) **18**, a motion estimation and compensation circuit (ME/MC) **19**, and a quantization adapter (QA) **20**.

Briefly summarized, the known encoder operates as follows. The input video frame X is divided into blocks of 8×8 pixels. The difference between each pixel block of input frame X and the corresponding block of a prediction frame X_p is discrete cosine transformed into a block of 8×8 coefficients. The coefficients are subsequently quantized, by which perceptually irrelevant picture details are irreversibly removed (lossy compression). The quantized coefficients are variable-length encoded and stored in a buffer from which the signal is applied to a transmission channel or record carrier. The encoded frame is locally decoded by inverse quantization, inverse discrete cosine transformation, and addition to the prediction frame X_p. The reconstructed frame is stored in the frame memory and subjected to motion estimation and compensation so as to constitute the prediction frame for the next input frame.

The encoder includes a quantization adapter **20** for calculating the quantization steps with which the DCT-coefficients are quantized. In this embodiment, the MPEG2 quantization mechanism is used in which a predetermined quantization matrix, which defines the step sizes to be applied to the respective coefficients of an 8×8 coefficient block, is multiplied by a quantization scale factor q (herein further referred to as quantization parameter). The quantization parameter is adapted from frame to frame, but may be 'modulated' within a frame as a function of local image details. The quantization parameter may be controlled to represent a given image quality (resulting in a variable bit rate) or a given bit rate (resulting in a variable quality). Various embodiments of quantization adapters (also referred to as bit rate controllers) are known in the art and may be employed in the encoder according to the invention.

The arrangement in accordance with the invention increases the quantization parameter q for selected frames, thereby degrading the image quality of said frames but reducing their bit costs. In this embodiment, the arrangement includes a multiplier **23** which multiplies the quantization parameter q calculated by the quantization adapter **20** by a predetermined factor F (e.g. F=1.4). A switch **22** has a position P in which the conventional quantization parameter q is applied to the quantizer **12** and a position P' in which the coarser quantization parameter F.q is applied to the quantizer. The switch is controlled by a control circuit **22** in a predetermined manner. For example, the control circuit selects every other P-frame to be more coarsely quantized.

Fig. 2A shows a diagram illustrating the performance of a conventional MPEG2 encoder which produces a stream of IPPP.. frames (no B-frames). Each frame is quantized in accordance with the quantization parameter q as calculated by the quantization

adapter **20**. The lower row of figures denotes the bit cost of the respective frame, expressed as a percentage of the bit cost of the respective I-frame. The bit cost of P-frames appears to be 38% in this example.

Fig. 2B shows a similar diagram for an encoder in accordance with the invention. The quantization adapter **20** has been set to produce the same image quality as in Fig. 2A. In accordance therewith, the bit cost for I-frames is the same as in Fig. 2A. Every other P-frame (denoted P' in the Figure) is now quantized with the quantization parameter 1.4q. The bit cost of the P'-frames is thereby reduced from 38% to 26%. The image quality of said frames is reduced in proportion therewith. The surprising effect of the invention is that the gain in bit cost is not lost in the subsequent 'conventional' P-frames. As shown in Fig. 2B, the bit cost of 'conventional' P-frames increases only from 38% to 42%. The net result is a considerable reduction of the bit rate at the same perceptual image quality (or a higher perceptual quality at the same bit rate) of the encoded video stream. In a practical experiment, the bit rate of a typical video signal was reduced from 15.2 Mbit/sec to 12.9 Mbit/sec at the same perceptual quality.

It is to be noted that the bit stream produced by an MPEG encoder in accordance with the invention fully complies with the MPEG standard. It should also be noted that although the invention has been described with reference to an IPP.. encoder (no B-frames), the invention does not exclude B-frame encoding. For example, an encoder may produce an IBPBP.. sequence in which selected P-frames have been quantized with the coarser quantization parameter. The coarser quantization parameter may even be applied to I-frames to the extent that such I-frames are used as prediction frames for subsequent frames.

Fig. 3 shows embodiments of arrangements for transmitting and receiving a video signal. An encoder 100 receives an image signal I on input 102. The encoder 100 is preferably constructed in accordance with the embodiment of Fig. 1, but may be any kind of compressed video signal encoder that supplies compressed video signals including first quantization parameters (q) representing a first quality or bit rate for quantizing selected first frames (P) of predictively encoded frames, and second quantization parameters (F.q) representing a second quality or bit rate that is lower than the first quality or bit rate for quantizing selected second frames (P') of the video signal. The compressed video signal is supplied at output 106 to a transmitter 108. The transmitter 108 then converts the compressed video signal into a transmission signal and applies the transmission signal to a transmission medium 110. The transmitter 108 and the transmission medium 110 may take any known form, for example, in the case of a broadcast, the transmitter 108 modulates the compressed

video signal onto a radio frequency (RF) carrier wave, and the transmission medium 110 may comprise airwaves containing the radio waves, or a cable carrying the radio waves. Alternatively, it is known to supply digital signals via the Internet. As such, the transmitter 108 may comprises means for transmitting the compressed video signal over the transmission medium 110 comprising the Internet.

A receiver 112 coupled to the transmission medium 110 then receives the transmission signal and applies the image signal to a display device 118.

Fig. 4 shows a block diagram of embodiments of arrangements for recording and playing back a compressed video signal onto and from a record carrier. This arrangement is substantially similar to that shown in Fig. 3, except that the compressed video signal at the output 106 of the encoder 100 is applied to a recording device 120. The recording device 120 then converts the compressed video signal into a recording signal and records the recording signal onto the record carrier 122. The recording device 120 and the record carrier 122 may take any known form. For example, in the case of the record carrier 122 being a magnetic video tape, the recording device 120 may take the form of a helical scan video tape recorder. Alternatively, the record carrier 122 may be an optical disk, for example, CD-ROM, CD-R, DVD, DVD-ROM, DVD-R/W, etc. In that event, the recording device 120 would take the form of an optical disk recorder.

In order to play back the recording signal RS, the record carrier 122 is inserted into a playback device 124 which processes the recording signal RS and generates an image signal. As with the embodiment of Fig. 3, the playback device 124 applies the image signal to the display device 118.

The invention can be summarized as follows. The concept of B-frames gives the MPEG video compression standard its high encoding efficiency. However, B-frame encoding roughly doubles the complexity of an MPEG encoder. In view thereof, MPEG encoders have been developed which produce I-frames and P-frames only. They are less complex but also less efficient. To improve the efficiency of such "IPP encoders", selected P-frames are quantized more coarsely than other P-frames, for example, by multiplying the conventional quantization step size by 1.4. Although this results in isolated frames ("virtual B-frames") being encoded with a lower quality, the overall perceptual quality is not affected. It has been found that the gain in bit rate obtained by the coarser quantization is not lost in subsequent P-frames, even though the subsequent frames are encoded with reference to the lower quality frames.

It should be noted that the above-mentioned embodiments illustrate rather than limit the invention, and that those skilled in the art will be able to design many alternative embodiments without departing from the scope of the appended claims. In the claims, any reference signs placed between parentheses shall not be construed as limiting the claim. The word "comprising" does not exclude the presence of elements or steps other than those listed in a claim. The word "a" or "an" preceding an element does not exclude the presence of a plurality of such elements. The invention can be implemented by means of hardware comprising several distinct elements, and by means of a suitably programmed computer. In the device claim enumerating several means, several of these means can be embodied by one and the same item of hardware. The mere fact that certain measures are recited in mutually different dependent claims does not indicate that a combination of these measures cannot be used to advantage.

CLAIMS:

1. A method of compressing a video signal, the method comprising:

predictively encoding (10,11) frames (X) of said video signal with reference to a prediction frame (X_p),

calculating (20) a quantization parameter (q) for each encoded frame,

quantizing (12) the encoded frames in accordance with said quantization parameter,

characterized in that said step of calculating the quantization parameter includes calculating a first quantization parameter (q) representing a first quality or bit rate for quantizing selected first frames (P) of said predictively encoded frames, and a second quantization parameter (F.q) representing a second quality or bit rate that is lower than said first quality or bit rate for quantizing selected second frames (P') of the video signal, the method further including:

decompressing (15-18) the compressed second frames to constitute the prediction frame (X_p) for predictively encoding the first frames.

2. A method as claimed in claim 1, wherein the step of calculating the second quantization parameter includes calculating said first quantization parameter (q) and multiplying (23) said first quantization parameter by a given factor (F).

3. A method as claimed in claim 1, wherein said predictively encoded frames constitute a series of successive frames, the second selected frames being every other frame of said series.

4. An arrangement for compressing a video signal, the arrangement comprising:

encoding means (10,11) for predictively encoding frames (X) of said video signal with reference to a prediction frame (X_p),

calculation means (20) for calculating a quantization parameter (q) for each encoded frame,

a quantizer (12) for quantizing the encoded frames in accordance with said quantization parameter,

characterized in that said calculation means (20) are arranged to calculate a first quantization parameter representing a first quality or bit rate for quantizing selected first frames (P) of said predictively encoded frames, and a second quantization parameter (F.q) representing a second quality or bit rate that is lower than said first quality or bit rate for quantizing selected second frames (P') of the video signal, the arrangement further including:

means (15-18) for decompressing the compressed second frames to constitute said prediction frame (X_p) for predictively encoding first selected frames.

5. An arrangement as claimed in claim 4, wherein said calculation means (20) comprise a multiplier (23) for multiplying the first quantization parameter (q) by a given factor (F).

6. An arrangement as claimed in claim 4, wherein said predictively encoded frames constitute a series of successive frames, the second selected frames being every other frame of said series.

7. A compressed video signal, comprising:

a prediction frame (X_p),

predictively encoded (10,11) frames (X) that have been predictively encoded with reference to the prediction frame (X_p),

respective quantization parameters (q) for respective encoded frames, the encoded frames having been quantized (12) in accordance with said respective quantization parameters, the quantization parameters including first quantization parameters (q) representing a first quality or bit rate for quantizing selected first frames (P) of said predictively encoded frames, and second quantization parameters (F.q) representing a second quality or bit rate that is lower than said first quality or bit rate for quantizing selected second frames (P') of the video signal.

8. A storage medium (122) on which the compressed video signal of claim 7 has been stored.

9. A method of transmitting or recording a video signal, the method comprising:

generating the compressed video signal of claim 7; and

transmitting or storing the compressed video signal.

10. An arrangement for transmitting or recording a video signal, the arrangement comprising:

means (100) for generating the compressed video signal of claim 7; and

means (108, 120) for transmitting or recording the compressed video signal.

1/2

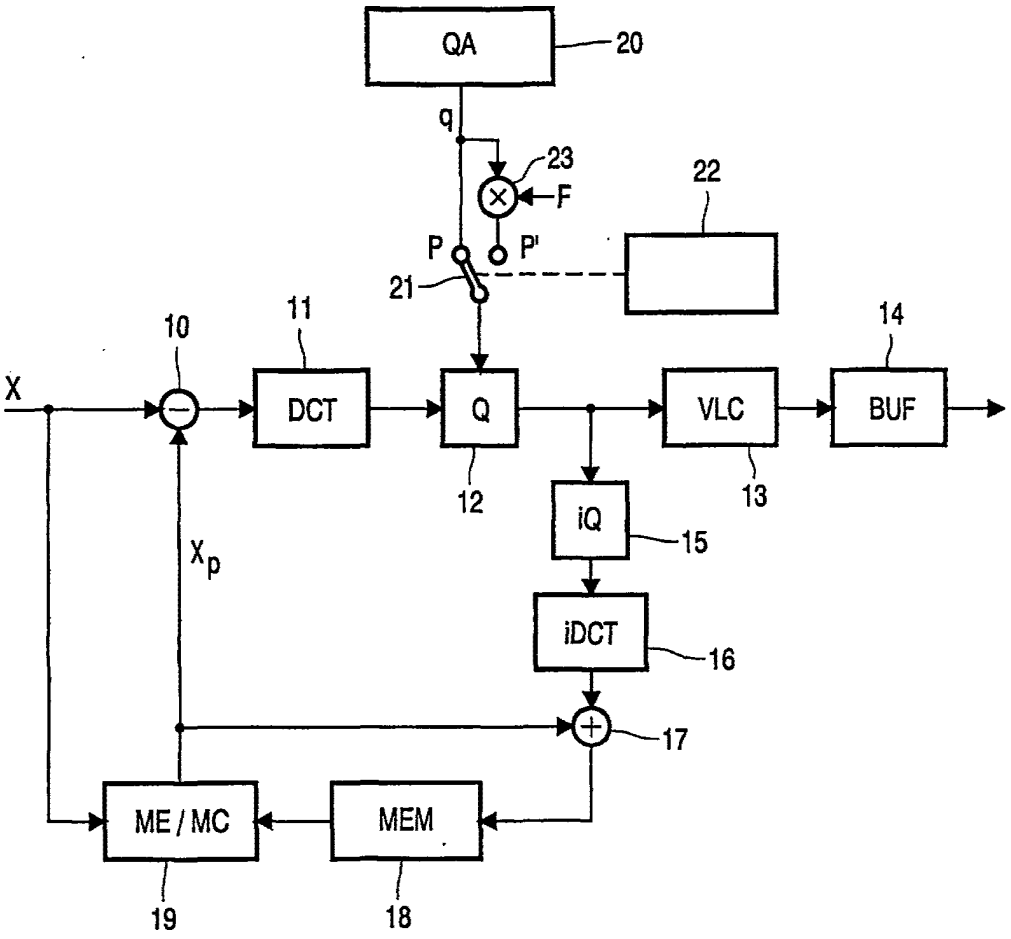


FIG. 1

I	P	P	P	P	P	I	P	P	P	P	P
100%	38%	38%	38%	38%	38%	100%	38%	38%	38%	38%	38%

FIG. 2A

I	P'	P	P'	P	P'	I	P'	P	P'	P	P'
100%	26%	42%	26%	42%	26%	100%	26%	42%	26%	42%	26%

FIG. 2B

2/2

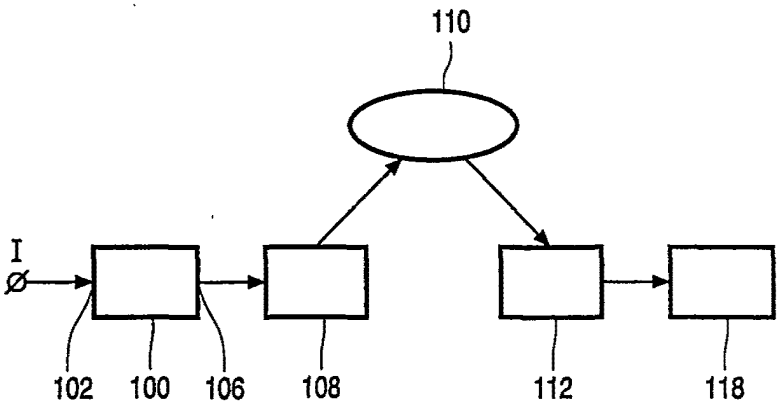


FIG. 3

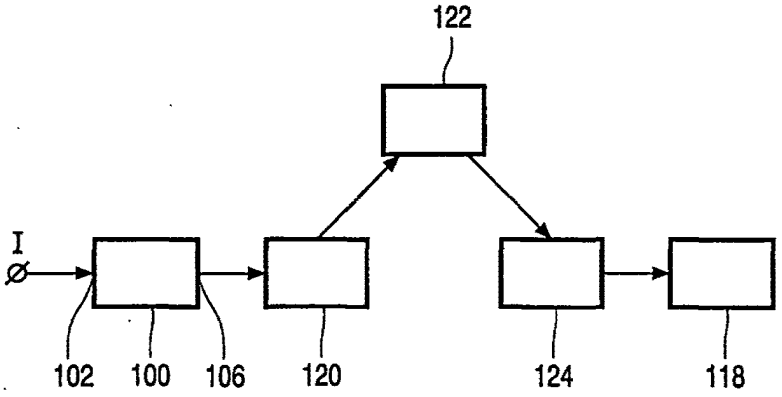


FIG. 4

INTERNATIONAL SEARCH REPORT

International Application No
I/EP 01/04319

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 H04N7/50

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 H04N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
INSPEC, EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WENGER S: "Temporal scalability using P-pictures for low-latency applications" 1998 IEEE SECOND WORKSHOP ON MULTIMEDIA SIGNAL PROCESSING (CAT.NO.98EX175), 1998 IEEE SECOND WORKSHOP ON MULTIMEDIA SIGNAL PROCESSING, 7 - 9 December 1998, pages 559-564, XP002177102 Redondo Beach, CA, USA the whole document	1-10
A	US 5 832 121 A (ANDO YUJI) 3 November 1998 (1998-11-03) column 11, line 64 -column 14, line 11	1,4,7,9, 10

☐ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
10 September 2001	24/09/2001

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Marie-Julie, J-M

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Information on patent family members

National Application No
T/EP 01/04319

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 5832121 A	03-11-1998	JP 8098185 A	12-04-1996
		EP 1098530 A	09-05-2001
		EP 0705041 A	03-04-1996
		US 5757968 A	26-05-1998

Form PCT/ISA/210 (patent family annex) (July 1992)

OIPE IAP78
JUN 16 2008
PATENT & TRADEMARK OFFICE

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant:	HANNUKSELA	Examiner:	Unassigned
Serial No.:	11/338,934	Group Art Unit:	2613
Filed:	January 25, 2006	Docket No.:	TPO.004.C2
Title:	GROUPING OF IMAGE FRAMES IN VIDEO CODING		

CERTIFICATE UNDER 37 CFR 1.8: The undersigned hereby certifies that this correspondence and the papers, as described hereinabove, are being deposited in the United States Postal Service, as first class mail, in an envelope addressed to: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450, on June 10, 2008.

By: Rennae Johnson

SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT (37 C.F.R. § 1.97(b))

Mail Stop Amendment
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

With regard to the above-identified application, the item of information listed on the enclosed Form 1449 is brought to the attention of the Examiner.

This statement should be considered because it is submitted before the mailing date of a first Office Action on the merits for the above-identified application. Accordingly, no fee is due for consideration of the item listed on the enclosed Form 1449.

This item was also first cited in a communication from a foreign patent office regarding a counterpart foreign application not more than three months prior to the filing of this information disclosure statement.

Applicant has not provided a copy of the U.S. publication/patent listed on the enclosed Form 1449 in accordance with 37 C.F.R. §1.98(a)(2).

No representation is made that a reference is "prior art" within the meaning of 35 U.S.C. §§ 102 and 103 and Applicant reserves the right, pursuant to 37 C.F.R. § 1.131 or otherwise, to

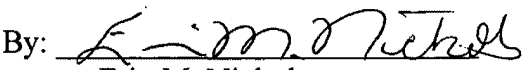
establish that the reference(s) are not "prior art." Moreover, Applicant does not represent that a reference has been thoroughly reviewed or that any relevance of any portion of a reference is intended.

Consideration of the items listed is respectfully requested. Pursuant to the provisions of MPEP § 609, it is requested that the Examiner return a copy of the attached Form 1449, marked as being considered and initialed by the Examiner, to the undersigned with the next official communication.

Respectfully submitted,

HOLLINGSWORTH & FUNK, LLC
8009 34th Avenue South, Suite 125
Minneapolis, MN 55425
952.854.2700

Date: June 10, 2008

By: 
Erin M. Nichols
Reg. No. 57,125

FORM 1449* INFORMATION DISCLOSURE STATEMENT IN AN APPLICATION (Use several sheets if necessary)	Docket Number: TPO.004.C2	Application Number: 11/338,934
	Applicant: HANNUKSELA	
	Filing Date: Jan. 25, 2006	Group Art Unit: 2613

O I P E IAP76 JUN 16 2008 PATENT & TRADEMARK OFFICE

U.S. PATENT DOCUMENTS

EXAMINER INITIAL	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	20010024472	09/27/2001	Sporer et al.			

FOREIGN PATENT DOCUMENTS

	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	TRANSLATION	
						YES	NO

OTHER DOCUMENTS (Including Author, Title, Date, Pertinent Pages, Etc.)

EXAMINER	DATE CONSIDERED
EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609; draw line through citation if not in conformance and not considered. Include copy of this form for next communication to the Applicant.	

*Substitute Disclosure Statement Form (PTO-1449) Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	PATENT NUMBER	GROUP ART UNIT	FILE WRAPPER LOCATION
11/338,934		2621	

OC000000032242965

Correspondence Address/Fee Address Change

The following fields have been set to Customer Number 76385 on 09/23/2008

- **Correspondence Address**
- **Power of Attorney Address**

The address of record for Customer Number 76385 is:

76385
Hollingsworth & Funk, LLC
8009 34th Avenue South
Suite 125
Minneapolis, MN 54425

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant:	HANNUKSELA	Examiner:	Unassigned
Serial No.:	11/338,934	Group Art Unit:	2613
Filed:	January 25, 2006	Docket No.:	TPO.004.C2
Confirmation No.:	6610	Customer No.:	76385
Title:	GROUPING OF IMAGE FRAMES IN VIDEO CODING		

CERTIFICATE UNDER 37 CFR 1.8: The undersigned hereby certifies that this correspondence is being transmitted to the U.S. Patent and Trademark Office via the Office electronic filing system in accordance with 37 CFR 1.6(a)(4) on March 12, 2009.

By: /Tracey M. Dotter/
Tracey M. Dotter

SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT (37 C.F.R. § 1.97(b))

Mail Stop Amendment
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

With regard to the above-identified application, the items of information listed on the enclosed Form 1449 are brought to the attention of the Examiner.

This statement should be considered because it is submitted before the mailing date of a first Office Action on the merits for the above-identified application. Accordingly, no fee is due for consideration of the items listed on the enclosed Form 1449.

These items were also first cited in a communication from a foreign patent office regarding a counterpart foreign application not more than three months prior to the filing of this information disclosure statement.

In accordance with 37 C.F.R. § 1.98(a)(2), copies of the foreign references listed on the enclosed Form 1449 are provided. As indicated in the communication from the Chinese patent office (a copy with an English translation is attached) the English-language EP publication corresponds to the cited Chinese reference.

Please note that any notations or markings on the attached documents do not reflect particular relevance, or lack thereof, to the present application, nor were they necessarily made by anyone affiliated with the prosecution of the present application.

No representation is made that a reference is "prior art" within the meaning of 35 U.S.C. §§ 102 and 103 and Applicant reserves the right, pursuant to 37 C.F.R. § 1.131 or otherwise, to establish that the reference(s) are not "prior art." Moreover, Applicant does not represent that a reference has been thoroughly reviewed or that any relevance of any portion of a reference is intended.

Consideration of the items listed is respectfully requested. Pursuant to the provisions of MPEP § 609, it is requested that the Examiner return a copy of the attached Form 1449, marked as being considered and initialed by the Examiner, to the undersigned with the next official communication.

Respectfully submitted,

HOLLINGSWORTH & FUNK, LLC
8009 34th Avenue South, Suite 125
Minneapolis, MN 55425
952.854.2700

Date: March 12, 2009

By: /Erin M. Nichols/
Erin M. Nichols
Reg. No. 57,125

FORM 1449* INFORMATION DISCLOSURE STATEMENT IN AN APPLICATION (Use several sheets if necessary)	Docket Number: TPO.004.C2	Application Number: 11/338,934
	Applicant: HANNUKSELA	
	Filing Date: Jan. 25, 2006	Group Art Unit: 2613

U.S. PATENT DOCUMENTS

EXAMINER INITIAL	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE

FOREIGN PATENT DOCUMENTS

	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	TRANSLATION	
						YES	NO
	EP 0798932	10/1/1997	EP			NA	
	CN 1173783A	2/18/1998	China				X

OTHER DOCUMENTS (Including Author, Title, Date, Pertinent Pages, Etc.)

EXAMINER	DATE CONSIDERED
EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609; draw line through citation if not in conformance and not considered. Include copy of this form for next communication to the Applicant.	

[19]中华人民共和国专利局

[51]Int.Cl[6]
H04N 7/06

[12] 发明专利申请公开说明书

[21] 申请号 97103739.6

[43]公开日 1998 年 2 月 18 日

[11] 公开号 CN 1173783A

[22]申请日 97.3.31

[30]优先权

[32]96.3.29 [33]US[31]08/626,184

[71]申请人 松下电器产业株式会社

地址 日本大阪

[72]发明人 塞卜瑞舍·V·莱姆波利
约瑟夫·P·奥哈拉
爱德温·罗伯特·梅依尔
罗伯特·T·莱恩

[74]专利代理机构 永新专利商标代理有限公司

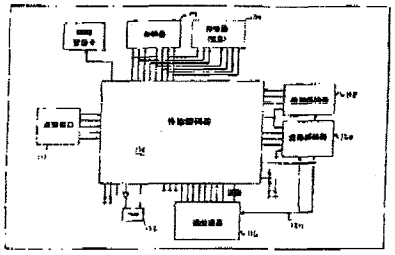
代理人 程 伟

权利要求书 3 页 说明书 24 页 附图页数 8 页

[54]发明名称 用于处理非连续程序数据的数字视频解码器

[57]摘要

接收和处理使用 MPEG-2 格式的传输数据流的传输解码器包括：与物理层通道接口的连接，缓冲存储器，主微处理器，音频和视频解码器，时钟信号电路。收到不连续性指示位时对传输包首部进行语法分析并将有效部分存于存储器。收到下一程序计数器参考（PCR）值时将其载入计数器。在存储器内搜索时间标记，若发出则将其恢复。在时间标记值前约一帧时刻设置定时器中断，该中断发生时，程序计数器参考值被送往解码器，为解码器提供足够数据处理时间。



(BJ)第 1456 号

权 利 要 求 书

1. 一种用于传输解码器系统的方法，所述系统用于对传输数据流解码，该数据流包括由首部和有效部分组成的传输包，首部包括不连续性指示位，所述系统包括至少一个缓冲区和至少一个解码器，前者用于临时存储接收到的传输数据流，后者用于对来自缓冲区的数据输出解码供显示之用，所述每个传输解码器和解码器都有各自的计数器，此方法用于在设定了不连续状态后处理程序计数器参考（PCR）的接收，其特征在于包括以下步骤：

对接收到的传输包进行语法分析，并将数据存入存储器；

一旦接收到下一个 PCR 值，将接收到的程序计数器参考值载入传输解码器计数器中；

对 PCR 值之后接收并储存在存储器中的数据进行数据标记搜索并将检索该数据标记值；

基于检索后的时间标记值设置定时器中断；以及

当定时器中断发生时，将 PCR 值传送给解码器的计数器，从而为解码器提供充足的时间处理数据。

2. 根据权利要求 1 所述的方法，其中，在时间标记值之前不足一帧的时间处设置定时器中断。

3. 根据权利要求 1 所述的方法，其中的搜索的步骤包括：

一旦接收到含有时间标记值的传输包，中断一台主微控制器以发出与存储的传输包相关的主微控制器地址信息；以及

接着，在存储的传输包的适当地址位置搜索时间标记值。

4. 根据权利要求 1 所述的方法，还包括以下步骤：

在进行语法分析之前，通过确定接收到的传输包的边界是否与预确定的包边界符合使它们生效，并且只对生效的包进行语法分析。

5. 根据权利要求 4 所述的方法，其中使接收到的传输包生效的步骤包括检查接收到传输包中的同步位。

6. 在用于接收传输数据流的系统中的一种处理不连续性指示位的接收的装置，该数据流包括由首部和有效部分组成的传输包，首部包括不连续性指示位，所述系统包括至少一个缓冲区和至少一个解码器，前者用于临时存储接收到的传输数据流，后者用于解码来自缓冲区的数据输出供显示之用，所述每个传输解码器和解码器都有分离的计数器，所述装置的特征在于包括：

接收和检测不连续性指示位装置；

对传输包进行语法分析并将数据存储在存储器中的装置；

一旦接收到下一个程序计数器参考值，将接收到的程序计数器参考值载入传输解码器计数器中的装置；

在存储器中存储的数据中搜索时间标记并在发现后检索时间标记值的装置；

基于时间标记值设置定时器中断的装置；以及

当定时器中断发生时，将程序计数器参考值发送到该计数器的装置。

7. 在一个用于解码传输数据流的传输解码器系统中，该数据流包括由首部和有效部分组成的传输包，首部包括不连续性指示位，所述传输解码器与至少一个外部解码器配合以解码来自传输解码器的数据输出，所述每个传输解码器和外部解码器都有分离的计数器以保持

与数据流同步，一种处理不连续性指示位的接收的方法包括以下步骤：

接收不连续性指示位；

对传输包进行语法分析并将数据存储在存储器中；

一旦接收到下一个程序计数器参考值，将接收到的程序计数器参考值载入传输解码器计数器中；

在存储器存储的数据中搜索时间标记并在发现后检索时间标记值；

在时间标记值前大概一帧的时间处设置定时器中断；以及

当定时器中断发生时，将程序计数器参考值发送到解码器的计数器中，从而为解码器提供充足的时间进行数据处理。

说 明 书

用于处理非连续程序数据的数字视频解码器

本发明涉及根据 MPEG-2 标准编码的数据进行数据接收、处理、及传输，具体来说，本发明涉及处理 MPEG-2 格式数据流的传输解码器的操作和接口。

高清晰度电视(HDTV)在试图替代传统的电视机方面继续取得进展。多家公司和协会正努力制定多种标准以供应 HDTV 的全球市场，从而为这种进展铺平道路。

其中有一个公司集团又被称为“数字高清晰度电视大联盟”，它包括AT&T、David Sarnoff研究中心、麻省工学院和其他一些机构。Robert Hopkins 在 1994 年夏季的《消费者电子 IEEE 交易》上发表了一篇以“北美数字陆地高清晰度电视：大联盟高清晰度电视系统”为名的文章，全面概述了该集团的长足进步。在此将引用该文章关于 HDTV 系统的背景和基础知识方面的教导以供参考，其中包括对程序和传输包流的使用。

大联盟采用由国际标准组织(ISO)下属的动态图像专家组(MPEG)开发的 MPEG-2 标准对视频和音频信息进行编码。被接受的标准定期发表在下列刊物上：动态图像和相关音频 ISO/IEC 13818 - 2(1995 年)的通用编码：信息工程的视频部分(下文简称为“视频部分”)，动态图像和相关音频 ISO/IEC 13818 - 1(1994 年 11 月)的信息工程的系统部分(下文简称为“系统部分”)。在此将引用二者关于已建立的标准和格式方面的教学资料以供参考。

MPEG-2 标准的语法定义了几种数据记录层，用来传送音频和视频数据。例如，为了传输信息，一个表示多视频序列的数字数据流被分为

多个较小的单元。每个单元被分别封装成多个打包的基本流(PES)包中。为了传输需要，每一个打包的基本流包被依次分割成多个定长传输包。每个传输包包含的数据只与一个打包的基本流包有关。传输包还包括一个首部，其中带有控制信息，有时也包括适应字段子段，用于对传输包解码。然而，打包的基本流数据也可能包括多种基本流。

正如在系统部分所述，多种基本流之间的同步是通过定期在数据流中提供时间标记实现的。例如，N 个基本流的解码是通过将流的解码校正到一个通用主时基的方式实现的，而不是调整一种流的解码以适应另一种流的解码。主时基可能是 N 个解码器中某一个的时钟信号，或者是数据源的时间信号，还可能是一个外部提供的时钟信号。

为了有助于该同步处理，MPEG-2 标准的传输首部中还包括一个不连续性指示位。系统部分提供了一种语法表示以显示一个典型的“不连续性指示位”在传输首部中的位置。该语法在以下的表 1 中有具体描述。

表 1

传输流适应字段

语法	位数	助记码
adaptation_field() {		
adaptation_field length	8	imsbf
if (adaptation_field_length > 0) {		
discontinuity_indicator	1	bslbf
random_access_indicator	1	bslbf
elementary_stream_priority_indicator	1	bslbf
PCR_flag	1	bslbf
OPCR_flag	1	bslbf
splicing_point_flag	1	bslbf
transport_private_data_flag	1	bslbf
adaptation_field_extension_flag	1	bslbf
if (PCR_flag == '1') {		
program_clock_reference_base	33	uimsbf

reserved	6	bslbf
program_clock_reference_extension	9	uimsbf
}		
if (OPCR_flag == '1') {		
original_program_clock_reference_base	33	uimsbf
reserved	6	bslbf
original_program_clock_reference_extension	9	uimsbf
}		
if (splicing_point_flag == '1') {		
splice_countdown	8	tcimsbf
}		
if (transport_private_data_flag == '1') {		
transport_private_data_length	8	uimsbf
for (i=0; i<transport_private_data_length; i++){		
private_data_byte	8	bslbf
}		
}		
if (adaptation_field_extension_flag == '1') {		
adaptation_field_extension_length	8	uimsbf
ltw_flag	1	bslbf
piecewise_rate_flag	1	bslbf
seamless_splice_flag	1	bslbf
reserved	5	bslbf
if(ltw_flag == '1'){		
ltw_valid_flag	1	bslbf
ltw_offset	15	uimsbf
}		
if(piecewise_rate_flag == '1'){		
reserved	2	bslbf
piecewise_rate	22	uimsbf
}		
if(seamless_splice_flag == '1'){		
splice_type	4	bslbf
DTS_next_au[32..30]	3	bslbf
marker_bit	1	bslbf
DTS_next_au[29..15]	15	bslbf

```
                    marker_bit                          1       bslbf
                    DTS_next_au[14..0]                  15      bslbf
                    marker_bit                          1       bslbf
                }
                for (i=0;i<N;i++){
                    reserved                            8       bslbf
                }
            }
            for (i=0);i<N;i++){
                stuffing_byte                           8       bslbf
            }
        }
    }
```

如系统部分中所述，不连续性指示位在传输首部内的适应字段中。不连续性指示位是一个 1 位的字段，设为逻辑值“ 1 ”时，表示当前的传输流包的不连续性状态为真。当不连续性指示位设为逻辑值“ 0 ”或不存在时，不连续性状态为假。不连续性指示位用于表示下列两种不连续性状态：系统时基不连续和连续性计数器不连续。

通过在被指定为程序时钟参考（ Program Clock Reference — PCR ）程序标识 的传输流程序标识（ PID ）中使用不连续性指示位，可以指示系统时基的不连续性（参见系统部分第 50 页的 2.4.4.9 ）。接下来，可以接收到一个程序时钟参考（ PCR ）值（即所谓主时钟信号）。

例如，当使用一个时基记录的节目材料被插入采用另一个时基记录的数据流中时，设定不连续性指示位。时基中的差别表现为一个不可知的程序时钟参考（ PCR ）值。

当一个被指定为 PCR 程序标识 的程序标识 的传输流包的不连续状态为真时，使用同一个程序标识 的传输流包中下一个 PCR 表示相关节目的新系统时间时钟取样。当包括一个新系统时基的 PCR 的包的第一个字

节到达系统目标解码器（ STD ）的输入端时，系统时基不连续点被定义为这个时刻。

值得注意的是：系统目标解码器在先有技术中广为人知，简而言之，它是表示一个在概念上与传输编码器的输出端相连的假定解码器。要求的条件是符合 MPEG-2 标准的数据流不会引起系统目标解码器的数据溢出问题（其中包括缓冲区溢出）。对系统目标解码器的其他讨论在系统部分给出。

紧接着，在传输流包中不连续指示位也可能被置为“ 1 ”。这些传输包都属于包含新的系统时基 PCR 的包的前一个程序时钟参考 程序标识。在这种情况下，一旦不连续性指示位被置为“ 1 ”，在同一个程序 PCR 标识的所有传输流包中它都被置“ 1 ”，该程序 PCR 标识符合和包括包含新系统时基的第一个 PCR 的传输流包。当系统时基的不连续性发生时，新的系统时基的不少于两个 PCR 在另一个系统时基不连续性发生之前被接收。进一步来说，除非欺骗模式状态为真，否则来自不多于两个系统时基中的数据会在任意时间出现在一个程序的系统目标解码器缓冲区中。

当时基不连续性要求更新多系统时间常数（ STC ）值时（即在传输解码器的时钟管理器和在外部的视频和音频解码器中），特别是当等待时间可能存在于正被外部解码器解码的数据和正被接收并储存在传输解码器的缓冲区中的数据之间时。换句话说，典型的情况是一个传输解码器与外部的视频和音频解码器耦合，后者中每一个都保存一个 STC 或等价的值。因此，当系统时基发生改变时，同时校正所有的 STC 值就可能引发系统的同步问题。

考虑到这个因素，本发明所要解决的一个重要问题在于传输解码器怎样才能有效地处理缓冲器和解码器之间的系统时基不连续性问题。

本发明包括一个用于接收传输数据流的系统，该数据流包括具有不连续性指示位的首部和有效内容的传输包。该系统至少包括一个缓冲存储器和一个外部解码器，前者用于临时储存接收到的传输数据流，后者用于解码来自缓冲区的数据输出以供显示之用。本发明用于处理对不连续性指示位的接收，其手段是通过接收到不连续性指示位后对传输包的首部进行语法分析，以及将传输包的有效内容存入存储器。一旦接收到下一个程序计数器参考值，一个计数器中就会装入接收到的程序计数器参考值。然后在储存的数据中搜寻时间标记，一旦发现，就检索该时间标记。最后，基于已检索到的时间标记为某一个时间点设置一个定时器中断，并且当定时器中断发生时，将程序计数器参考值送往解码器，同时向解码器提供新时间标记值不充分时间，使之能处理含有新的时基的数据，而无需干扰对含有前一时基的数据的处理。

图 1A 示出了一个典型的数字传输解码器及其各种接口的高级功能框图。

图 1B 示出了一个高级流程图，描述了采用图 1A 所示的实施例时本发明的一个方面采取的步骤。

图 2A 示出了图 1A 所示的传输解码器的典型实现的高级功能框图。

图 2B 示出了图 1A 所示的传输解码器的典型实现的高级数据/控制流程框图。

图 3 示出了图 2 所示的传输解码器中使用的存储器控制器的功能框图。

图 4 示出了用以描述存储器控制器如何处理指针 FIFO 项的典型流程图。

图 5A 示出了在本发明的视频解码器接口部分执行的步骤的典型流程图。

图 5B 示出了在本发明的音频解码器接口部分执行的步骤的典型流程图。

概述

如上文所述，本发明的一个重要方面在于解决传输解码器如何高效地处理不连续性指示位的问题。本发明的另一个重要方面在于处理传输解码器和仅能处理基本流数据的视频解码器之间的接口问题。另外，本发明还有一个方面涉及传输解码器和位速率限制音频解码器之间的接口问题。为了全面介绍以上的各个方面，本发明包括一个特别设计的 MPEG-2 传输解码器。

如果一个节目使用 MPEG-2 格式通过卫星向全国的多个接收站播放，一个时基不连续的实例就出现了。这就是为介绍本发明的不连续性部分提供的具体环境。但是，在节目播出的过程中，每个接收站会加入自己的商业广告节目。在这种情况下，应设置一个不连续性位来指示节目时基中的空白或间断，以便于自己拥有本地时基的商业广告的插入。

然而，当各个接收站插入各自的节目时，来自卫星的节目传输并不会中断。由此可知，当遇到不连续性指示位时，接收来自本地接收站并带有插入广告内容的节目的传输解码器的内部定时会跟踪适当的时基。

再看看附图，图 1A 示出了一个典型的数字传输解码器 110 的高级功能框图以及各种接口。如图 1A 所示，传输解码器 110 包括与以下各部分的连接：物理层通道接口（又叫通道接口） 112 、缓冲存储器 114 、主微处理器 116 、外部视频和音频解码器 118 和 120 以及时钟信号电路（例如 VCXO ） 122 。在本发明的典型实施例中，传输解码器 110 支持一个视频接口和多达两个音频接口（例如，音频接口 A 和 B ）。

处于传输解码器 110 外部的主微处理器 116 也通过微总线 24 与视频和音频解码器 118 和 120 耦合。这种连接可以使用并行或串行数据路径，并根据被选择使用的单个外部解码器的要求而变化。在任何情况下，将主微处理器 116 和外部解码器连接起来的方法都是众所周知的，同时也为拥有普通技艺的人员所理解。

一般来说，传输解码器 110 和每个外部视频和音频解码器 118 和 120 都包括一个的系统时间常数（ STC ）值（图中未给出）。该值储存在寄存器或计数器中用于使接收到的数据流的解码和显示同步。

如上所述，当时基不连续在数据流中发生时可能出现问题，而且这种不连续要求 STC 值在传输解码器 110 和视频及音频解码器 118 和 120 中同时更新。特别容易发生问题的时间是当等待时间可能存在于正被外部解码器解码的数据和正被接收并储存在传输处理器 110 中的数据之间时。

在本发明中，微处理器 116 在控制数据流的同步和视频及音频解码器 118 和 120 的运作方面起着十分重要的作用。在本发明中，微处理器 116对传输解码器110中的多个芯片寄存器能进行直接的存储器映射式存取。微处理器 116 还能通过读、写、和水印指针对外部存储器 114 内的缓冲区和芯片上 32 位可读写 FIFO （如图 3 中所示）进行存取。微处理器 116 通过一个内部微总线 223（如图 2B 所示）与多个内部功能块连接，从而允许各种事件和错误状态通过可屏蔽中断用信号发出。

一般来说，如图 1B 所示的流程图中，本发明的一个方面通过特别设计的传输解码器 110 在接收到不连续性指示位（步骤 150 ）后执行以下步骤：

1) 由于不连续性指示位不必和下一个程序时钟参考值（ PCR)处于同一个数据包中，所以当下一个程序时钟参考 值被接收后，位于传输

解码器 110 中的时钟管理电路（如图 2A 中所示）内的计数器中将装载下一个程序时钟参考值（ PCR ）（步骤 152 ）；

3) 在经过语法分析并存储在存储器中的后续包中搜索时间标记，一旦找到某一个，则将检索时间标记值（步骤 154 ）；

4) 由微处理器 116 在某个时间点设置定时器中断，该时间点只比检索的时间标记值提前不到一帧的时间（步骤 156 ）；而且

5) 当定时器中断发生时，当前 STC 值被送往外部视频和音频解码器 118 和 120 （步骤 158 ），从而给外部解码器足够的时间以抢在数据流中的时间标记值与保存在视频或音频解码器 118 或 120 中的计数器值进行比较之前处理数据。

下文将详细介绍这些高级步骤的执行情况。

图 2A 和图 2B 示出了图 1A 所示的传输解码器 110 的典型执行过程的高级功能框图。下文参照图 2A 和 2B ，介绍传输解码器 110 中的总体数据流动情况。

如图所示，图 1A 的通道接口 112 与通道数据处理器（ CDP ） 210 相连。通道接口 112 还向 NRSS 接口 212 提供数据，以便用 NRSS 智能卡 230 加密。在此特结合 NRSS 委员会（ 4/95 ）的条件存取 EIA 标准 2.6 版关于 NRSS 智能卡的论述以为参考。

通道数据处理器 210 被提供用于预处理（如检测、同步、确认等）所接收数据流中的包。处理完毕后，通道数据处理器 210 将输出数据发送到传输处理器（ TPROC ） 214 中供进一步处理之用。传输处理器 214 根据下文详细描述的程序标识表进行多种处理，例如语法分析等。

传输处理器 214 与传输解码器 110 中的多个功能块进行交互。这些功能块包括: 存储器控制器 216、时钟管理器 218，个人数据处理器 220 和高速数据端口 222 。如上所述，微处理器 116 能对多个芯片寄存器直接存取。这种存取是通过微接口 217 和内部微总线 223 （如图 2B 所示）获得的。存储器控制器 216 除了与传输处理器 214 连接外，还与外部存储器 114 和视频及音频解码器（ 224 、 226 和 228 ）相连。

现在介绍一下与单个功能块有关的数字数据流的处理细节。

通道数据处理器（ CDP ） 210

值得指出的是来自通道接口 112 中的数据流经过可能的缓冲后（图中未给出）被送往通道数据处理器 210 。

通道数据处理器 210 对接收到的数据流执行多种预处理步骤。具体来说，通道数据处理器 210 决定数据流中字节和帧的边界，并将接收到的数据流转换为 8 位的并行格式。在本发明的典型实施例中，在通道数据处理器 210 后在传输解码器 110 中进行的操作使用 字节格式进行。

为了确定字节和帧的边界，通道数据处理器 210 与传输包首部包含的帧模式同步。具体来说，这种同步是通过搜索 MPEG-2 的传输流同步字节（ sync_byte ）（例如: 47H ）并确认它在可编程数目的连续传输包开始处发生而实现的。

在此过程中，通道数据处理器 （ 210 ）核对同步字节，将数据流锁定并继续确认包。当同步位的搜索在进行中时，在 CDP 状态寄存器中设置搜索状态位（ Search_State ）（图中未给出）。一旦发现了同步字节并且对连续发生的确认在进行之中，同步状态位（ Sync_State ）就设定了。最后，当符合预先确定的搜索标准时，设定锁定状态位（ Lock_State ）。在本发明的典型实施例中，搜索标准是三个连续的确认过的包。同样，指示带有损坏同步位的包数目的同步滞后值被设定。对通道数据处理器

210 编程使这些同步位在宣布损失锁定状态之前就能通过。当这个滞后值被超过时，锁定的损失被宣布。通道数据处理器 210 然后发出一个锁定损失中断并且从锁定状态过渡回搜索状态。

另外，如果通道接口 112（如图 1A 所示）提供一个 CHPSTRT 信号，传输解码器 110 可能被设置以使用此信号作为一种手段来与数据流同步，方法是将通道数据处理器帧和同步控制寄存器中的帧模式位（ Framing_Mode ）设为逻辑值“ 1 ”。在此情况下，当 CHPSTRT 被确定后，在一个传输包时间内可获得同步。

另一种情况是，如果通道数据处理器帧和同步控制寄存器中的忽略同步位被置为逻辑值“ 0 ”，则通道数据处理器 210 在相对于 CHPSTRT 信号过渡过程的预定字节（第一个字节）处校验同步位的完整性，并当校验失败时宣布锁定状态。如果忽略同步（ Ignore-Sync ）设置为“ 1 ”，则通道数据处理器 210 不校验同步的完整性。

取得了锁定状态后，通道数据处理器 210 的输出被传输到传输处理器 214 并由其进行处理，后者以字节的格式对数据进行处理（例如，过滤、语法分析、格式化等）。

传输处理器（ TPROC ） 214

由传输处理器 214 产生的信号控制数据流动去往：存储器控制器 216 、微接口 217 、时钟管理器 218 、个人数据处理器 220 、和高速数据端口 220 。

在本发明的典型实施例中，传输处理器 214 在程序标识表的控制下处理数据流。如系统部分所述，一个程序标识是传输流首部中一个 13 位的字段，表明存储在包有效数据中的数据类型。一些程序标识值被分配，另一些则被保留。程序标识表的另一些细节在上文提到的 MPEG-2 规范中的系统部分的 3.3 节给出。

根据程序标识表，传输处理器 214 能同时处理多达 32 个用户可选择的程序标识。传输解码器 110 所处理的程序标识在程序标识表中均有规定。程序标识表中的位字段指定了各种数据处理选项。

在操作过程中，程序标识由传输处理器 214 从刚刚到达的传输包的首部中抽取出来，并同时与程序标识表中的所有项目比较。一旦发生匹配，就根据匹配程序标识表项中设定的选项对包中的数据进行处理。如果程序标识表项没有匹配的，就放弃该包。

有效部分格式位没有设定为放弃（ 000 ）的程序标识表中每一个地址都与一个通道号相连。存储器控制器 216 与传输处理器 214 相连，负责向 RAM 中特定的缓冲区（通道）传输数据。关于存储器控制器 216 的操作问题在下文中与图 2 、 3 、 4 有关的部分具体讨论。

在本发明的典型实施例中，保存于传输处理器 214 中的程序标识表由微处理器 116 在初始化过程中通过微总线 223 利用软件设置。

关于程序标识表项的格式，除了其他信息外，一个程序标识表项还包括： 1) 一个 13 位的程序标识， 和 2) 一个 1 位的 PCR 程序标识字段。后者一旦设定，就表明对于被解码的程序来说，此程序标识带有 PCR 。一个典型的程序标识表项还包括一个 1 位的打包的基本流 HDR 位，它一旦设定，就使得与打包的基本流首部中的“流标识”对应的 RAM 缓冲区地址储存于指针 FIFO 中。当存储器控制器 216 向指针 FIFO 中写入数据时，它发出一个中断。这个中断又被称为一个 DMA_MARK 信号，用于提醒微处理器 116 发送流标识（ Stream_ID ）字节。在本发明的典型实施例中，以基于表中位置的数字顺序向程序标识表项分配通道号。例如，向第一项分配 DMA 通道 0 ，向下一项分配 DMA 通道 1 等等。

传输处理器 214 表明带有程序时钟参考程序标识的传输包应该由时钟管理器 218 通过在程序标识表项中放入 PCR_PID 信号进行处理。时钟管理器（218）监视带有 PCR_flag（标志）的包的适应字段，并抽取 PCR。

传输处理器 214 根据包的程序标识将它们分开，然后除去传输包的首部，并结合存储器控制器 216 将程序标识包的有效部分（例如，打包的基本流包的部分）分别存储在外部存储器 114 中指定通道的顺序排列存储区字段。存储器 114 中的数据表示打包的基本流包。

传输处理器 214 还处理从主微处理器 116 向外部存储器进行的写操作。

存储器控制器 216

存储器控制器 216 将传输包中经过语法分析的有效部分存储在外部存储器 114 中，并根据需要向视频处理器 224、多达两个音频处理器 226 和 228 以及主微处理器 116 或微控制器提供视频、音频、PSI 和其他数据（如个人数据）。

图 3 给出了适用于图 2A 和图 2B 中显示的传输解码器 110 的存储器控制器的典型功能框图。如图 3 中所示，存储器控制器 216 通过 FIFO310 从传输处理器 214 中获得数据。存储器控制器 216 通过由微输入 FIFO314 和微输出 FIFO316 组成的微接口 217 从微处理器 116 中取得和传输数据。值得指出的是，图 2A 和图 2B 的存储器控制器 216 外部的微接口 217 在图 3 中也给出了，以便清楚地表示它与 FIFO314 和 316 的相互关系。

存储器控制器 216 还包括一个流程控制单元 318，它包括芯片存储器、一个 DMA 控制器 320、数据路径（Datapath）逻辑 322，以及 用于向视频和音频解码器（224、226 和 228）分别发送数据的。

存储器控制器 216 将来自下列两个来源中的数据写入外部存储器：其一为传输处理器 214（通过FIFO310 中的传输端口），其二为微处理器 116（通过 FIFO314 中的微端口）。

存储器控制器 216 从外部存储器中读取数据，并将其提供给以下四个部件：1) 视频接口（通过视频 FIFO324），2) 音频接口 A（通过音频接口 A 的 FIFO326），3) 音频接口 B（通过音频接口 B 的 FIFO328），4) 微控制器 116（通过微输出 FIFO316）。

外部存储空间被划分为多个分离的、互不重叠的缓冲区。其中每一个供一个 DMA 通道使用。可使用的通道多达 32 个之多。通道号与写入数据两个 FIFO 源或从四个目标 FIFO 中读出的数据相关连。

由传输处理器 214 提供并写入传入端口 FIFO310 的数据带有一个通道号标记。其它 FIFO 中每一个的通道号由微处理器 116 在 DMA 接口寄存器中定义。微处理器 116 为每一个被使用的通道设置一个 128 位的 DMA 通道配置寄存器（通道寄存器）。此通道寄存器含有由起始和结束地址定义的缓冲区空间等信息，而不管该缓冲区被设为 FIFO（读写指针互相缠绕，有时又叫环形队列）还是队列（ Queue ）（读写指针不会缠绕在一起），等等。

如果通道寄存器允许，DMA控制器320能在当前使用的缓冲区已满时获得一个新的缓冲区。这种特性通常只适用于那些存储为主微处理器 116 所用的数据的缓冲区，如 PSI 区字段。这些新缓冲区的起始和结束地址储存在新的缓冲 FIFO 中。只有微处理器 116 能向此 FIFO 中写入新的缓冲区。已使用过的缓冲区指针和其他类型的指针由 DMA 控制器 320 写入指针 FIFO 中。在本发明的典型实施例中，只有微处理器 116 能读取指针 FIFO。这些指针被微处理器 116 用于管理对缓冲区数据的读取和向新的缓冲 FIFO 返回使用过的缓冲区。图 4 中的指针 FIFO 项流程图给出了

在所有可能的信号条件、通道寄存器设置和缓冲区条件下指针 FIFO 的项。

在接下去的图 3 中，DMA 控制器 320 和数据路径 322 为外部存储器提供所有的接口信号。流程控制单元 318 通过向 DMA 控制器 320 提供适当的控制信号，使其为六个特定的接口 FIFO 中的某一个服务。这些信号包括:

1. DMA 控制器开始信号:

本信号命令 DMA 控制器 320 开始为使用诸如 DMA 控制器操作、DMA 输入 FIFO 选择和 DMA 输出 FIFO 选择之类命令的特定 FIFO 服务。同时，流程控制单元 318 读出 128 位特定通道寄存器的内容，并使其在通道寄存器（内部存储器）的数据总线上可用。DMA 控制器 320 将此 128 位的字读入它的内部寄存器，然后流程控制单元 318 被释放以根据内部存储器的需要执行其他的读/写操作。

2. DMA 控制器操作信号（1 位）:

该位发出一个读或写的信号，由接口 FIFO 上的 DMA 控制器执行。

3. DMA 输入 FIFO 选择信号（1 位）:

发出信号选择两个输入 FIFO（310 或 314）中的某一个提供服务。

4. DMA 输出 FIFO 选择信号（2 位）:

发出信号选择四个输出 FIFO（316、324、326 或 328）中的某一个提供服务。

5. 在请求/通道数据/通道号/DMA 控制器中的通道数据停止:

一般来说，DMA 控制器 320 持续为 FIFO 服务，直到因如接口 FIFO 满或外部存储器中无数据等原因不能继续服务为止。一旦这种情况发

生， DMA 控制器 320 分配请求的通道数据给流程控制单元 318 ，并且将通道寄存器数据和通道号数据提供给流程控制单元 318 。除非特定的 FIFO 需要急救行动（例如，因为它正接近满或空的状态），流程控制单元 318 使用通道数据更新通道寄存器存储器（例如，被更新的信息包括读写指针等），并指示 DMA 控制器 320 为回路中的下一个 FIFO 服务。

有些时候，流程控制单元 318 可能需要向 DMA 控制器 320 发送一个 DMA 控制器停止命令，作为对向另一个 FIFO 提供服务的需要作出的响应。该 FIFO 可能是拥有更高优先级的输入 FIFO 。这个命令指导 DMA 控制器 320 停止为当前正被其服务的 FIFO 服务，并将通道数据写入流程控制单元 318 。

DMA 控制器 320 使用以下信息更新通道寄存器：读指针、写指针和 DMA 控制器 320 停止为 FIFO 服务处的其他信息。由于流程控制单元 318 将这些信息写入了内部通道寄存器存储器， DMA 控制器 320 可以从中断点开始恢复为被中断的 FIFO 提供服务。

考虑到 DMA 控制器参与本发明不连续状态部分的整体运作过程，当包括一个流标识的字节的下一个 PES 包被 TPROC 214 写入存储器控制器 216 时，流程控制单元 318 引起 DMA_MARK 信号变高。 DMA_MARK 信号的过渡触发微处理器 116 处的中断，并且打包的基本流 包首部的地址被存储在微处理器已存取的输入指针 FIFO 中。由于这个打包的基本流首部可能有一个时间标记，所以对由不连续性指示位产生的标志位作出响应的微处理器 116 在储存的打包的基本流 首部中搜索时间标记。一旦发现，会检索此时间标记值。微处理器 116 基于检索到的时间标记值设置定时器中断。

再参看图 3 ，在本发明的典型实施例中，流程控制器单元 318 的内部存储器包括四个 64 × 32 位的块，每个块都分配给通道寄存器、新缓冲 FIFO 和指针 FIFO 。

在本发明的典型实施例中，每一个通道寄存器有 128 位并且最多有 32 个通道。所以有多达 128 × 32 个位以多个 128 位的格式分配给通道寄存器。

32 位的字被写入新缓冲区和指针 FIFO 或从它们中读出。至少 128 × 32 位的内存（即 128 个字）可用于新缓冲区和指针 FIFO 中。另外，如果使用的通道少于 32 个，通道寄存器中未使用的空间被分配给新的缓冲区和指针 FIFO 。

以下是 32 个 DMA 通道和 16 个 DMA 通道的存储器分配示例。

1) 32 个被使用的通道的存储器分配

新缓冲区 FIFO = 64 × 32 位

指针 FIFO = 64 × 32 位

通道寄存器 = 128 × 32 位

2)16 个被使用的通道的存储器分配

新缓冲区 FIFO = 64 × 32 位

指针 FIFO = 128 × 32 位

通道寄存器 = 128 × 16 位

时钟管理器 218

图 2A 和 2B 中所示的传输解码器 110 中的时钟管理器 218 使用连续 PCR 值同步本地 27MHz 的信号（由时钟信号电路 122 产生）。时钟管理器 218 中的计数器值被称为系统时间常数（ STC ）值。在本发明的典型实施例中，微处理器 116 通过内部微总线 223 存取计数器中的 STC 值。

时钟管理器 218 中的计数器定期与传输数据流中的 PCR 进行同步。当遇到数据流中的第一个 PCR 时，该 PCR 被“挤载”入计数器并成为初始 STC 值。当后来的 PCR 值被接收时，将它们与 STC 值相比较，以校

正时钟信号电路 122（例如，VCXO）的频率。在本发明中，此过程是通过使用在下文中会详细介绍的数字模拟转换器（DAC）实现的。一旦进行了相位错误校正，通常是将新的 PCR 值作为 STC 值载入。

一般每帧产生一次 PCR 值。在 MPEG-2 的标准中，至少每三帧内（大概十分之一秒）会产生一次 PCR 值。

典型的情况是，外部视频和音频解码器（118 和 120）通过更新外部解码器中的 STC 值与传输解码器 110 保持同步，该过程的实现基于时钟管理器 218 中保存的 STC 值。视频和音频解码器（118 和 120）通过将数据和接收到的打包的基本流 数据流中的时间标记相比较，使用储存值来确定何时解码或显示数据。值得指出的是，在本发明的典型实施例中，传输解码器 110、视频解码器 118 和音频解码器 120 都接收时钟管理器（218）产生的同一个 27MHz 的时钟信号。

传输解码器 218 包含对产生本地时钟信号的硬件支持。该信号对于某一个选定的程序来说是锁定在系统时钟信号上的。如上文所述，在特定的 PCR_PID 上接收的 PCR 值在串行接口上可为其他设备所用。

时钟管理器（218）包括一个系统时间计数器（STC）（图中未给出）、一个系统时间时钟寄存器（图中未给出）、一个程序时钟参考（PCR — Program Clock Reference）寄存器（图中未给出）、一个当前 STC 寄存器（图中未给出）、一个报警时钟中断寄存器（图中未给出）以及一个锁存从与选定的程序标识（PID）有关的首部中抽取的 PCR 值的逻辑部件（图中未给出）。它还包括串行 PCR 接口和用于控制外部系统时钟信号控制回路的 DAC 寄存器。

在本发明的典型实施例中，STC 是由 9 位基部和 33 位扩展部组成的 42 位计数器。

STC 基部将名义上的 27MHz 系统时钟输入信号分为 300 份。结果得到的 90KHz 信号驱动 STC 扩展。在任何时候， STC 都能被在微处理器 116 上运行的软件写或读。当遇到一个不连续性指示位后，可以将接收到的 PCR 值自动加载到 STC 中。

如上文所述，可以通过在程序标识表项中设定 PCR_PID 位，从刚到达的位流中自动抽取 PCR。处理接收到的 PCR 取决于软件设定的选项和接收 PCR 时当前传输包的不连续状态是否存在。

当包含 PCR_PID 的传输包被接收时，和传输首部中的不连续指示位为逻辑值 1 ，可以确认是不连续状态。该不连续状态将一直持续下去，直到包含 PCR_PID 的传输包被接收时不连续性指示位为逻辑值 0，或者当在 PCR_PID 包中发现下一个 PCR 为止。

如果设置为有效，当 PCR 被接收且不连续状态存在时，时钟管理器 218 自动将接收到的 PCR 载入 STC 中。此时，一个中断被发送到主微处理器 116。

时钟管理器 218 包括一个初始化位。该位被软件置为 1 时，则无论不连续状态存在与否，下一个被接收到的 PCR 都将被载入 STC 中。然后初始化位被时钟管理器 218 自动清除。只有当 STC 不包含有效的 PCR 时（即，当选择了新的程序标识时），才设置初始化位。

如果没有不连续状态存在或自动 STC 载入功能被置为无效，接收到的 PCR 被储存在 PCR 寄存器中，并且可选择向主微处理器 116 发出一个中断。在获得 PCR 的同时， STC 的当前状态被保存在 STC 寄存器中。储存下来的 PCR 和 STC 值被用于计算错误信号，从而控制外部系统的时钟信号（ XOCLK ）。

当接收到的传输包包括一个与 PCR_PID 等同的 PID（程序标识）和一个 PCR 时，此 PCR 被锁存在 PCR 寄存器中。此时 STC 计数器的当前

内容也被锁存在 STC 寄存器中，并且向主微处理器 116 发出一个中断。将两个值都锁存起来就消除了由于中断服务例程中存在变化的等待时间而产生的不确定性。在本发明的典型实施例中，两个寄存器都包括一个 9 位的基部和一个 33 位的扩展部。基部在 300 模块中，扩展部是用二进制表示的。在寄存器的内容被用于任何计算之前，它们可以被转化成二进制数。转化方法是将扩展部乘以 300，然后将与基部相加。

当前的 STC 寄存器是一个只读寄存器，它允许微处理器 116 在任何时间获得 STC 的当前内容。通过外部微总线 124，微处理器 116 能用这个值初始化视频解码器或音频解码器的 STC 。

如上文所述，传输解码器 110 包括一个 10 位的数字模拟转换器（ DAC ） 230 ，它在时钟管理器 218 的时钟校正控制回路中使用。DAC230 是只读的，也就是说，软件在存储器中存有最后载入 DAC 的值的副本。在运作中， DAC230 被用于为外部谐振晶体电压控制振荡器 VCXO122 产生控制电压。后者是系统时钟频率（名义上是 27MHz ）的来源。通过比较锁存的 PCR 和 STC 值，软件能确定本地时钟信号是否需要校正。例如，软件能获取 PCR 和 STC 之间的差别，并将这个错误条件加到 DAC230 的输入寄存器的当前内容中去。

总体来说，本发明的不连续部分包括以下内容:

1) 通道数据处理器（ CDP ） 210 核查接收到的传输包中的同步字节，以使该包生效。传输处理器（ TPROC ） 214 对传输首部进行语法分析，并将包的有效内容传送到存储器中的适当位置。根据传输首部中的程序标识值，数据被发送到存储器中的不同 FIFO 缓冲区（通道）中。

2) 因为接收到的不连续性指示位不必和下一个 PCR 值在同一个包中，所以接收到下一个 PCR 值时， TPROC （ 214 ）将“挤载”时钟管理器（ 218 ）中的计数器。

3) 当包括一个流标识字节的下一个打包的基本流包被传输处理器 214 写入存储器控制器 216 中时， DMA_MARK 信号变高。这就导致存储器控制器 216 向微处理器 116 发出中断并将打包的基本流 包首部的地址储存在一个输入指针 FIFO 中。这个打包的基本流 首部可能有一个时间标记（可能是解码时间标记或显示时间标记）。对来自存储器控制器 216 的中断作出响应的微处理器 116 在打包的基本流 首部中搜索时间标记，一旦发现，就检索时间标记值。

4) 然后，微处理器 116 在时间标记值前略短于一帧的时间处设置一个定时器中断。

5) 当此定时器中断发生时，微处理器 116 通过外部微总线 124 向视频和音频解码器发送新的 PCR 值，从而给了解码器足够的时间处理数据。这种处理发生在时间标记与视频或音频解码器中的计数器值比较之前，并且不会干扰与前一个 PCR 值相参照的数据的处理。

视频/音频解码器 224 、 226 、 228

如上文所述，一旦传输包被接收、处理和存储起来，存储器控制器 216 从存储器 114 中读出打包的基本流包，并将它们发送到外部视频和音频解码器中。然而，在打包的基本流 包到达外部解码器（ 118 和 120 ）前，它们被内部视频和音频处理器（ 224 、 226 和 228 ）处理。

这些内部处理器的一个重要功能在于确保离开传输解码器的数据流与外部解码器兼容。也就是说，不同的外部解码器有不同的数据需求。

例如，一些现在市场上可见的外部视频解码器能接受包括首部和有效部分的打包的基本流包流。这些流的形式主要与它被读出存储器时相同。另一方面，有些外部解码器只能接受对基本流（例如，打包的基本流 包的有效部分而非首部部分）的处理要求。另外，后一种类型的解码器除了要求有数据流外，还要求相关的控制信息，如时间标记和字节计

数等。由 C - Cube 制造的 CL9100 就是这样一种视频解码器。在此特参考其规格说明——CL9100 多模式视频解码器用户手册（94 年 10 月）——加以介绍。由此，视频处理器 224 可通过微处理器 116 进行设置，以格式化输出的数据流，使之与选定的外部解码器的要求兼容。

在处理 C - Cube CL9100 等外部视频解码器的接口方面，视频处理器 224 对打包的基本流 包的流进行处理，从而去除打包的基本流 的首部并将时间标记信息从打包的基本流 首部中取出来。另外，视频处理器 224 通过使用计数器（图中未给出），对发送给外部解码器的打包的基本流包保持字节计数。那么，视频处理器 224 包含时间标记信息，保持字节计数并将打包的基本流 包的有效部分传输给外部视频解码器。为了获得传送给外部解码器的时间标记和字节计数信息，视频处理器 224 向微处理器 116 发出一个中断。后者进行响应时，会通过内部微总线 223 从视频处理器 224 中读取时间标记和字节计数信息，并通过外部微总线 124 将这个控制信息发送到外部视频解码器中。

然而，有时在微处理器 116 能对来自视频处理器 224 的中断作出响应之前，另一个带有时间标记信息的打包的基本流 首部就到达了。例如，当微处理器 116 在很长的时间里被更高优先级的任务占用时，这种情况就可能发生。在这种情况下，视频处理器 224 将时间标记信息和状态一起储存起来。具体来说，视频处理器 224 会保存一个状态寄存器，表明微处理器是否“丢失”了时间标记。在本发明的典型实施例中，微处理器 116 最后才开始读取时间标记信息，如果它象状态中表示的那样丢失了一个或多个时间标记，那么它将只能恢复最新的时间标记。

另外，既然时间标记信息和字节计数信息将被提供给外部解码器，那么当微处理器 116 丢失了一个时间标记时，视频处理器 224 保持一种累积式的字节计数，因为许多数据标记可能在微处理器 116 读取信息之前就到达了。然后，微处理器 116 读取最新的时间标记和累积式的字节

计数,并将这些信息发送到外部视频解码器中。接下来，状态寄存器和字节计数器可以被重置。

如图 5A 中的流程图所示，接受到的数据的首部被分离并进行语法分析（步骤 510 ）。时间标记信息被存储并且字节计数被保存（步骤 510 ），由视频处理器 224 向微处理器 116 发出一个中断，提醒它时间标记信息可供读取（步骤 514 ）。如果附加的时间标记信息在微处理器读取前一个时间标记信息之前就已收到，新时间标记信息被存储，字节计数器继续积累（步骤 516 ），视频处理器 224 向微处理器 116 发出另一个中断，否则，微处理器读取状态寄存器、时间标记信息和字节计数信息并将它发送到外部视频解码器中（步骤 518 ）。

调节某个外部解码器的另一个例子是使用位速率限制音频解码器，例如 Zoran ZR38500 AC3 音频解码器。在此特引用其技术说明书（ 1994 年 10 月）以供参考。一般说来，大多数音频解码器提供一个数据请求，表明它们能接受音频数据，然后传输解码器 110 通过音频处理器（ 226/228 ）提供音频数据，直到外部音频处理器取消此请求。换句话说，这里存在着一种信号交换机制。

然而，位速率限制解码器，例如 Zoran AC3 音频解码器，要求音频 AC3 数据能作为音频数据帧发送，每一帧在首部中都包括一个帧时间和位速率。以高于预先确定的位速率发送数据可能会引起内部缓冲区溢出，从而导致数据丢失。进一步说，这些位速率限制解码器并不提供数据就绪的信号以表明它们已经准备好接收附加的音频数据。

所以，和视频解码器 224 一样，音频处理器 226 和 228 可被设置以调节选定的外部音频解码器。值得指出的是：在本发明的典型实施例中，除非是对于微处理器 116 可以存取的地址，音频处理器（ 226/228 ）在本质上是相同的。

因此，当音频处理器（226/228）被编程用于与该类音频解码器接口时，它们假定解码器一直在发送数据请求信号。然而，当打包的基本流包音频数据从存储器控制器216中提供给音频处理器（226/228）时，接收到数据的打包的基本流 首部被音频处理器进行语法分析。音频数据的帧大小和位速率被包含在经分析过的首部中。根据帧的大小和位速率，音频处理器（226/228）能确定帧边界（即帧时间）。有了帧时间信息，音频处理器（226/228）设置帧定时器，使得音频数据可用时，定时器被设定，一帧数据被触发到外部音频解码器及音频处理器（226/228）中，然后暂停，直到帧定时器超过时效。当定时器超过时效时，循环继续下去，每次一个音频帧，而另外的帧可用于向外部音频解码器中输送附加帧。控制信息，例如时间标记，由微处理器恢复并通过外部微总线发送出去。

如图 5B 中的流程图所示，从存储器中恢复打包的基本流包（步骤542），去除接收到数据的首部并进行语法分析（步骤530），确定一个帧时间(步骤532),定时器开始工作(步骤534),音频处理器（226/228）触发出一帧数据（步骤536），音频处理器（226/228）暂停，直到定时器超过时效（步骤538），尽管如图中所示，步骤542、530和532能在暂停期间执行。然后，此过程延续下去。

尽管在上述特定的实施例中，对跟踪 MPEG-2 传输数据流不连续性的系统时间常数进行更新的方法和设备进行了描述。但本发明并不局限于所示的细节。另外，无需脱离权利要求的范畴和本发明的精神实质，可实现细节的多种变化。

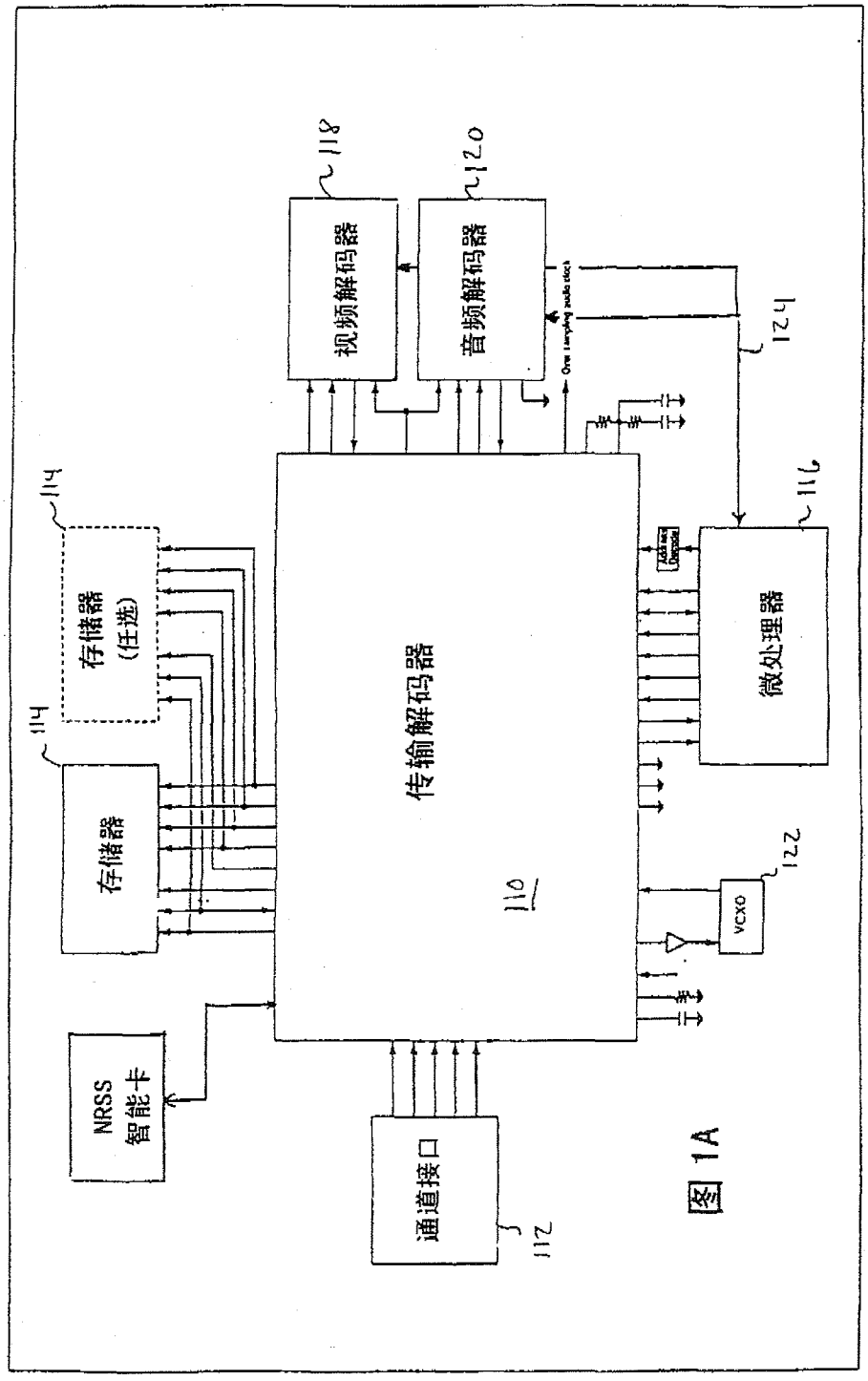

视频解码器
118
音频解码器
120
124
传输解码器
110
微处理器
116
存储器
(任选)
114
存储器
114
VCXO
122
NRSS
智能卡
通道接口
112

图 1A

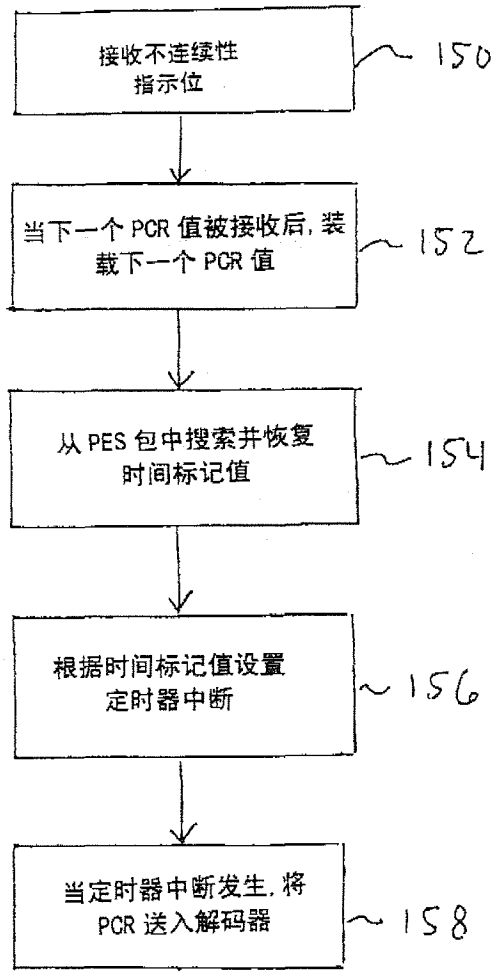

接收不连续性
指示位
150
当下一个 PCR 值被接收后，装
载下一个 PCR 值
152
从 PES 包中搜索并恢复
时间标记值
154
根据时间标记值设置
定时器中断
156
当定时器中断发生，将
PCR 送入解码器
158

图 1B

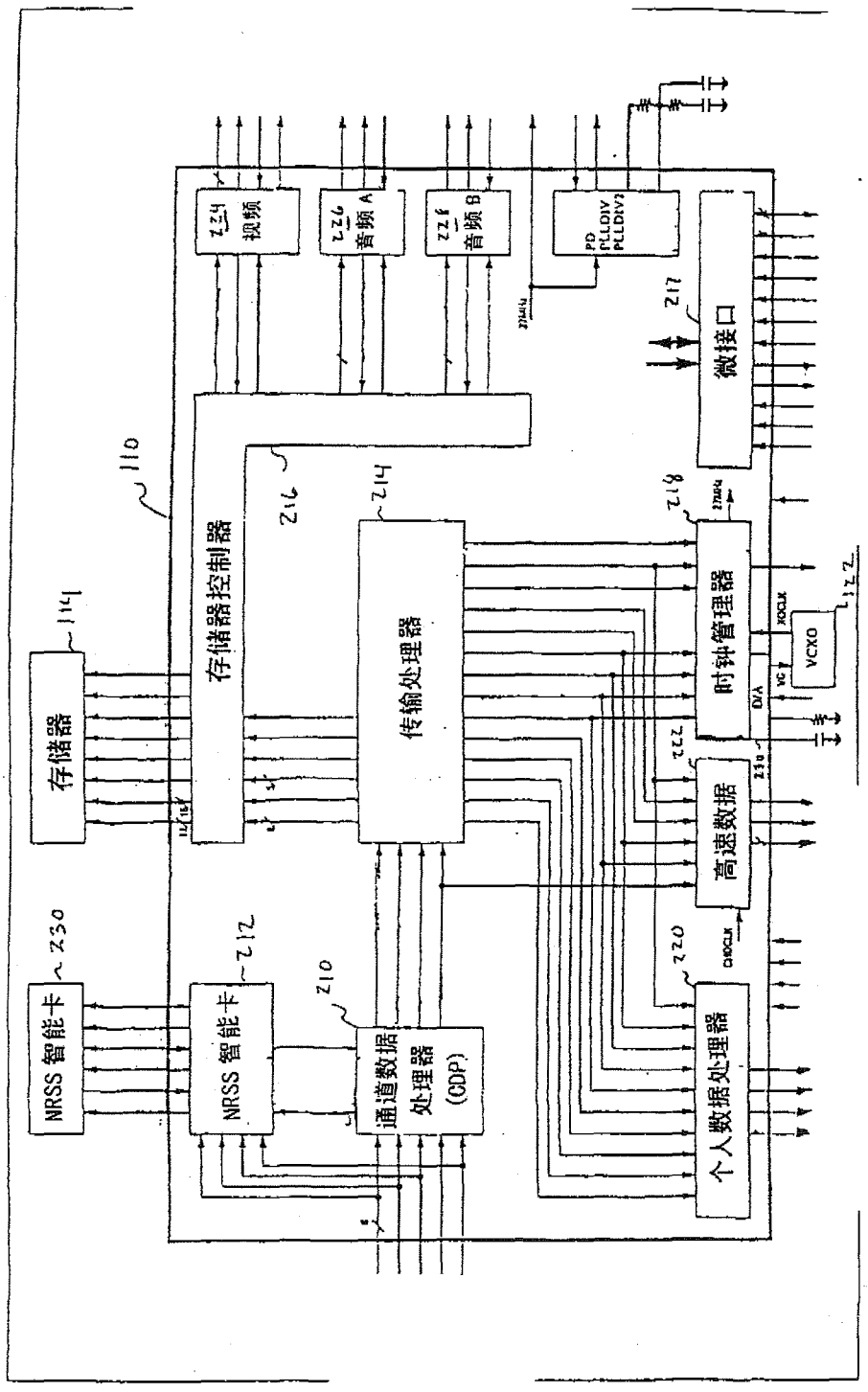

110
存储器
存储器控制器
传输处理器
时钟管理器
微接口
高速数据
个人数据处理器
通道数据处理器 (CDP)
NRSS 智能卡
NRSS 智能卡
视频
音频 A
音频 B
VCXO

图 2A

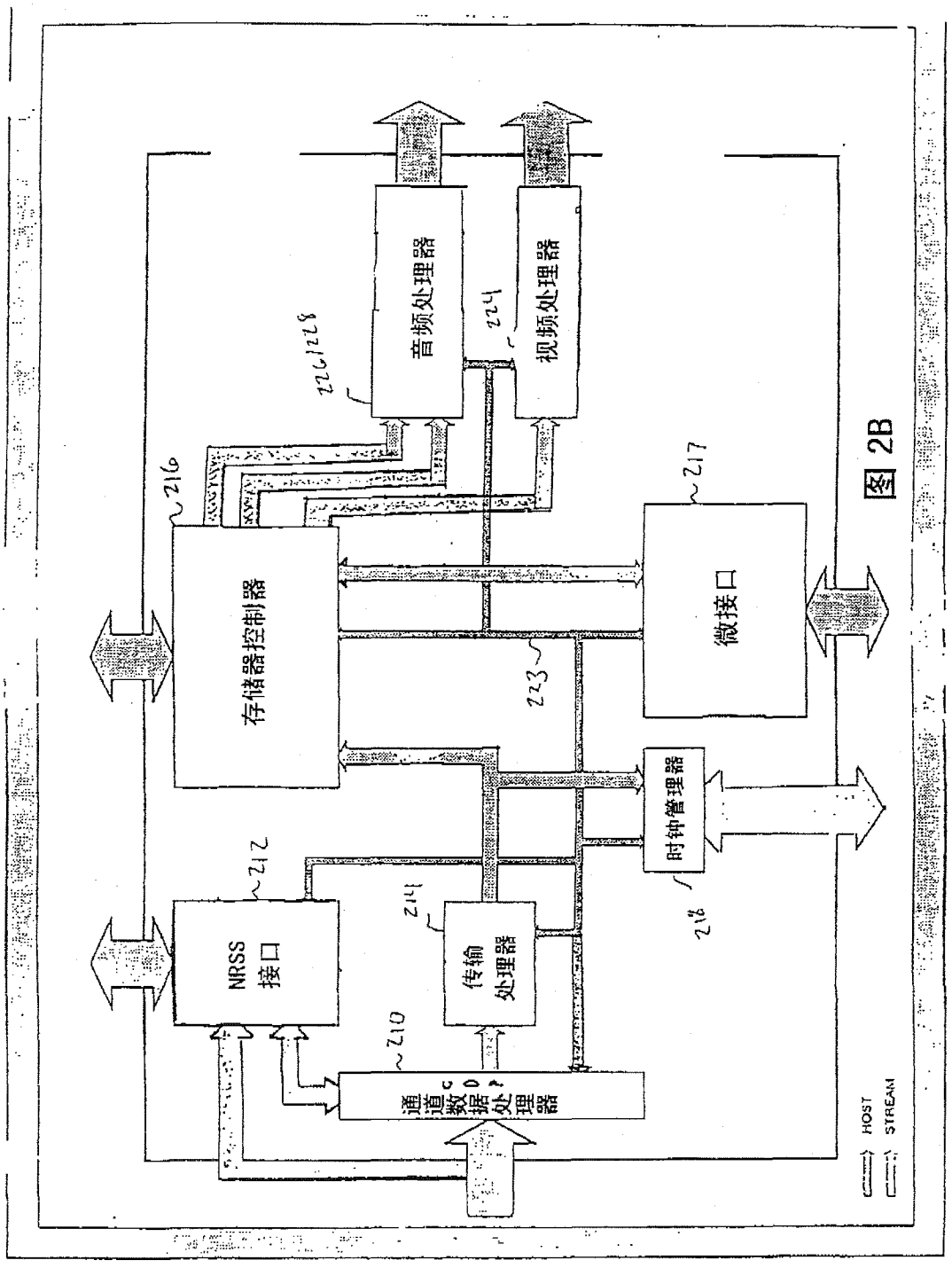

音频处理器
视频处理器
存储器控制器
微接口
NRSS
接口
传输
处理器
时钟管理器
通道数据处理器
HOST
STREAM

图 2B

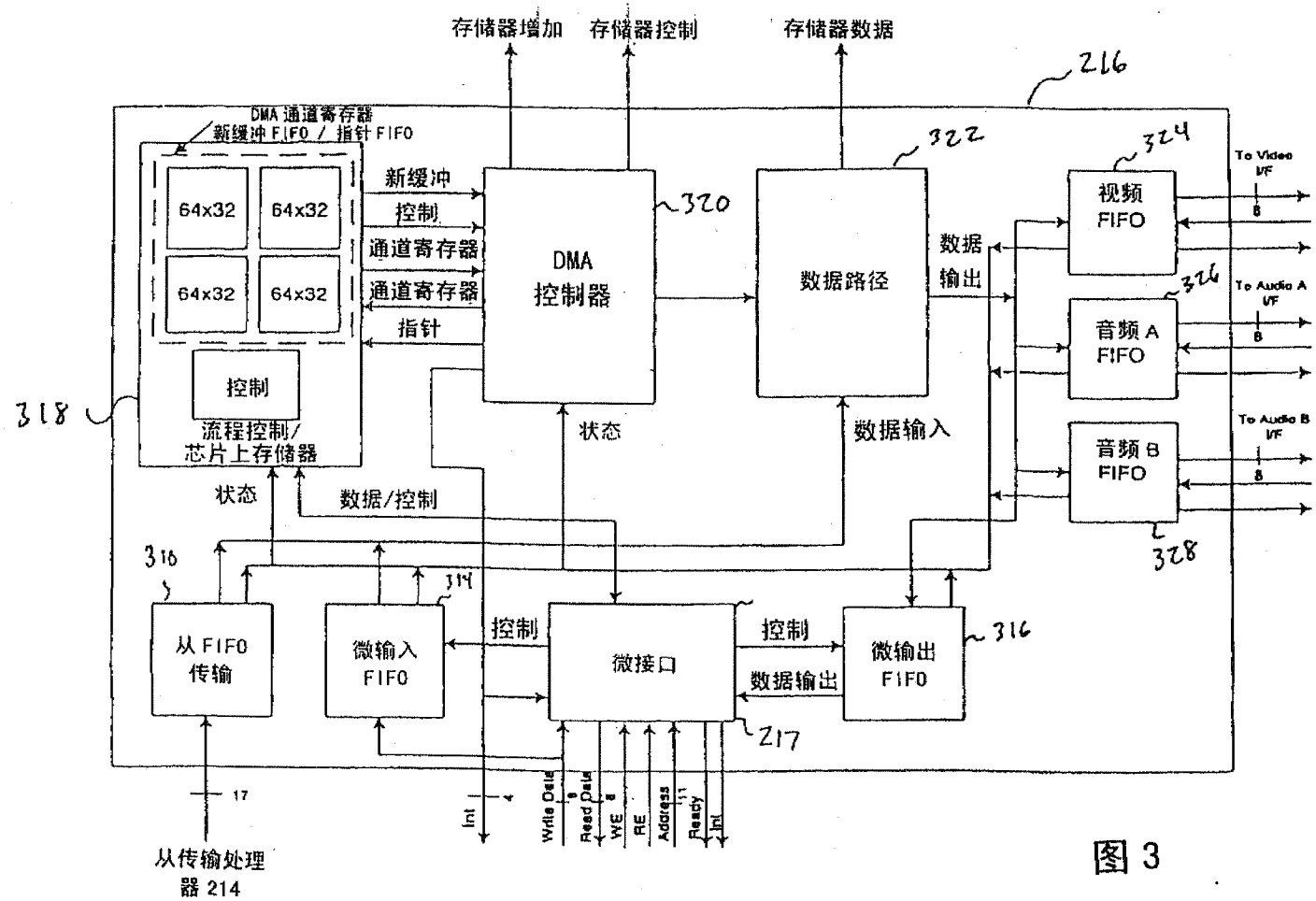
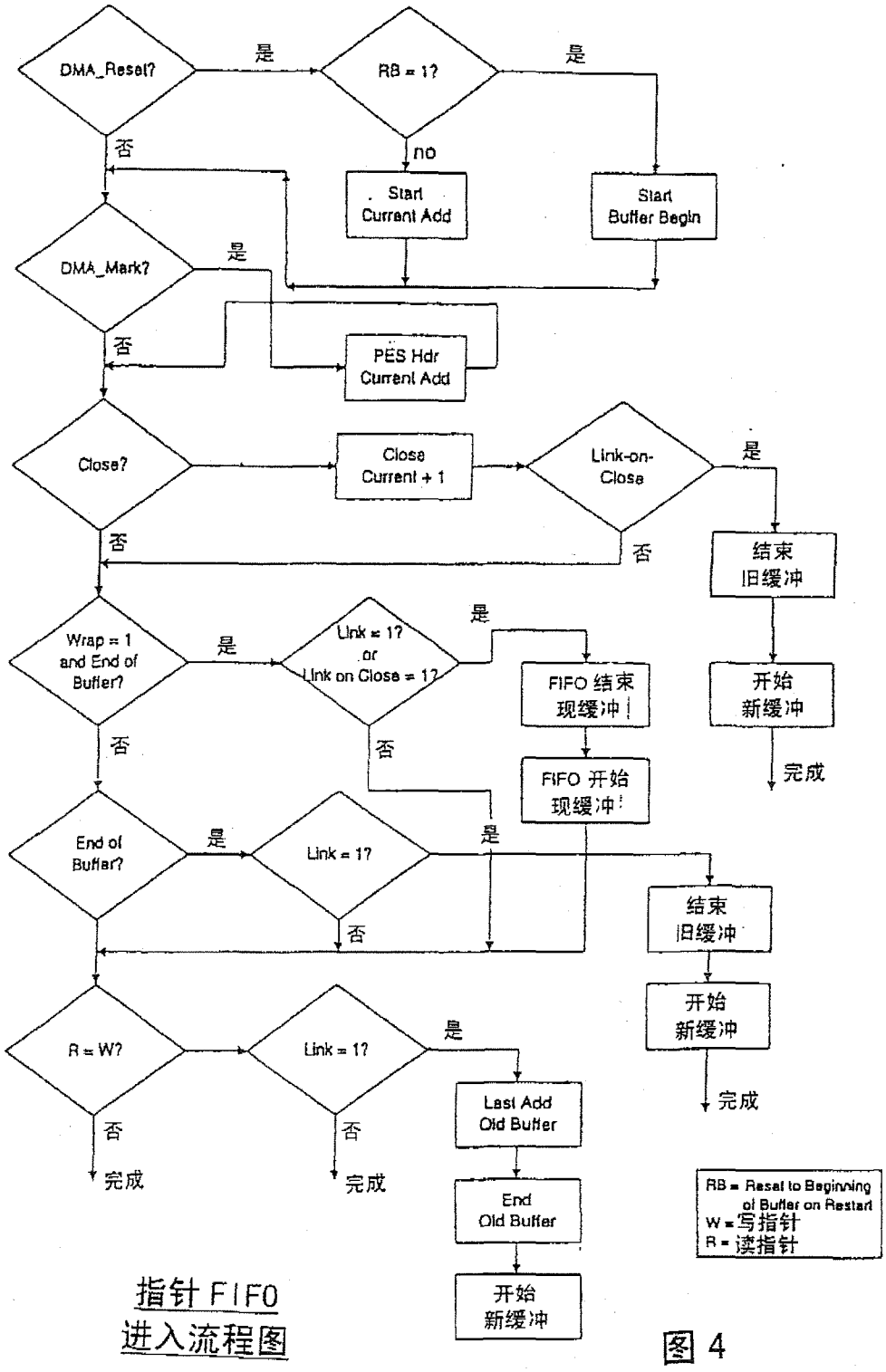

存储器增加
存储器控制
存储器数据
216
DMA 通道寄存器
新缓冲 FIFO / 指针 FIFO
64x32
64x32
64x32
64x32
控制
流程控制/
芯片上存储器
318
新缓冲
控制
通道寄存器
通道寄存器
指针
DMA
控制器
320
322
数据路径
数据
输出
视频
FIFO
324
音频 A
FIFO
326
音频 B
FIFO
328
To Video
I/F
To Audio A
I/F
To Audio B
I/F
状态
数据输入
状态
数据/控制
310
314
从 FIFO
传输
微输入
FIFO
控制
微接口
217
控制
数据输出
微输出
FIFO
316
17
从传输处理
器 214
Int
Write Data
Read Data
WE
RE
Address
Ready
Int

图 3

存储器控制器锁定



指针 FIFO 进入流程图

图 4

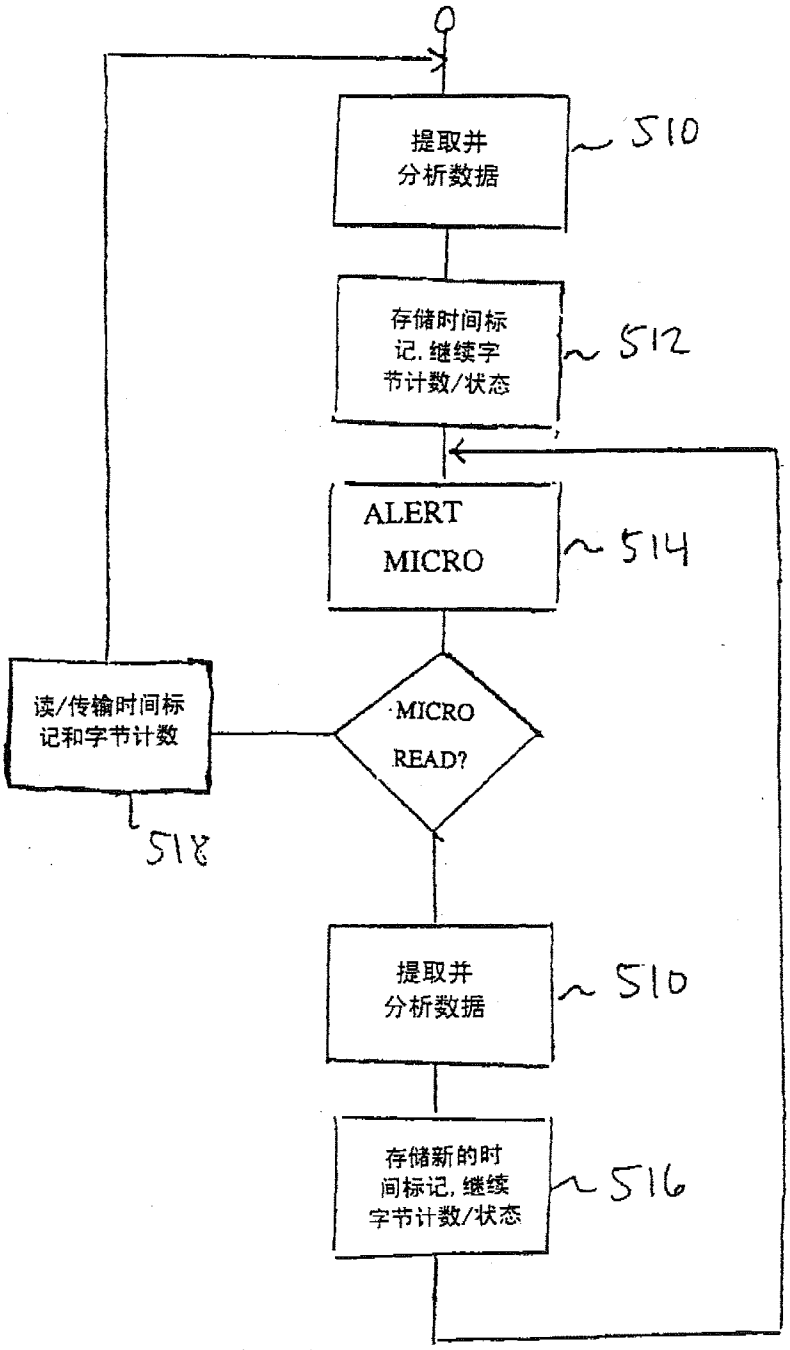

提取并
分析数据
510
存储时间标
记. 继续字
节计数/状态
512
ALERT
MICRO
514
MICRO
READ?
读/传输时间标
记和字节计数
518
提取并
分析数据
510
存储新的时
间标记, 继续
字节计数/状态
516

图 5A

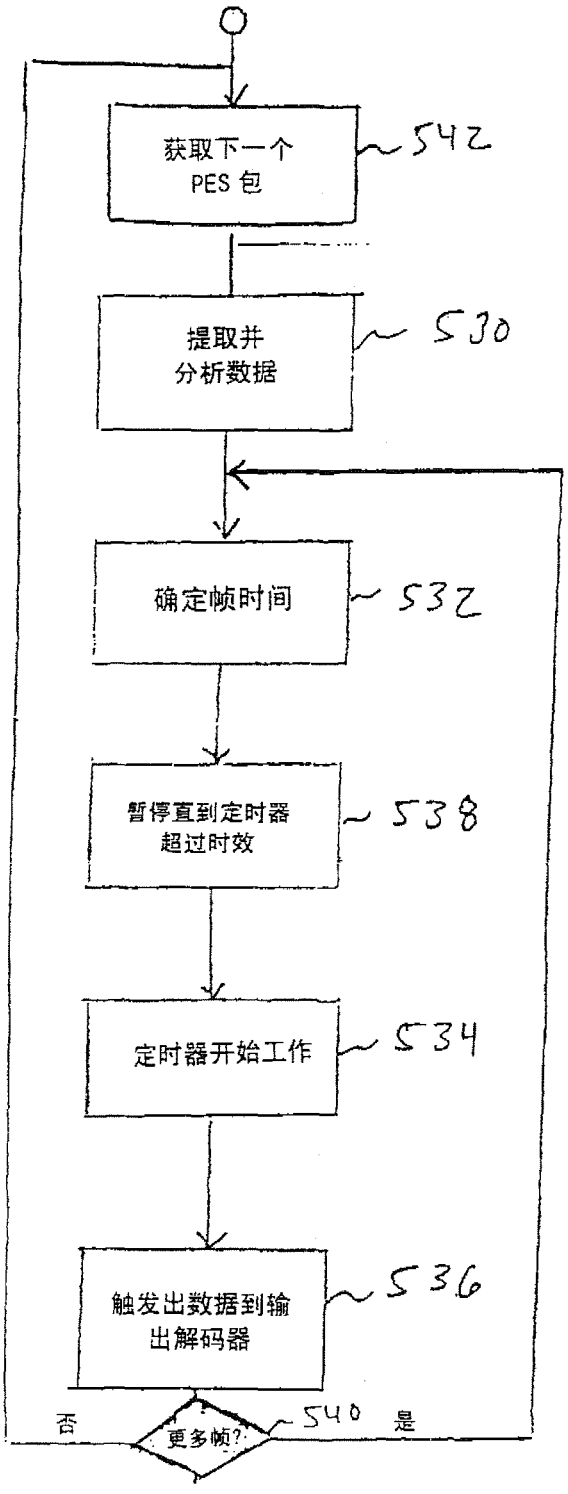

获取下一个
PES 包
542
提取并
分析数据
530
确定帧时间
532
暂停直到定时器
超过时效
538
定时器开始工作
534
触发出数据到输
出解码器
536
更多帧?
540
否
是

图 5B

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 0 798 932 A2

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
01.10.1997 Bulletin 1997/40

(51) Int. Cl.[6]: **H04N 7/50**

(21) Application number: **97104977.0**

(22) Date of filing: **24.03.1997**

(84) Designated Contracting States:
DE FR GB

(30) Priority: **29.03.1996 US 626184**

(71) Applicant:
MATSUSHITA ELECTRIC INDUSTRIAL CO., LTD.
Kadoma-shi, Osaka 571 (JP)

(72) Inventors:
- **Naimpally, Saipraisad V.**
  Langhorne, PA 19047 (US)
- **O'Hara, Joseph P.**
  Medford, NJ 08055 (US)
- **Meyer, Edwin Robert**
  Bensalem, PA 19020 (US)
- **Ryan, Robert T.**
  Langhorne,PA 19047 (US)

(74) Representative: **Schwabe - Sandmair - Marx**
Stuntzstrasse 16
81677 München (DE)

(54) MPEG-2 transport decoder

(57) A transport decoder 110, for receiving and processing a transport data stream using MPEG-2 formats, includes connections to a physical layer channel interface (channel interface) 112, a buffer memory 114, a host microprocessor 116, audio and video decoders 118/120, and clock signal circuitry 122. Upon receipt of a discontinuity indicator, the transport packet headers are parsed and transport packet payloads are stored in memory. Subsequently, upon receipt of a next program counter reference (PCR) value, a counter is loaded with the received program counter reference value. Next, the data stored in memory is searched for a time stamp and, when one is found, the time stamp is retrieved. Finally, a timer interrupt is set for a point in time approximately one frame time before the time stamp value, and, when the timer interrupt occurs, the program counter reference value is sent to the decoders, thereby providing sufficient time for the decoders to process data.

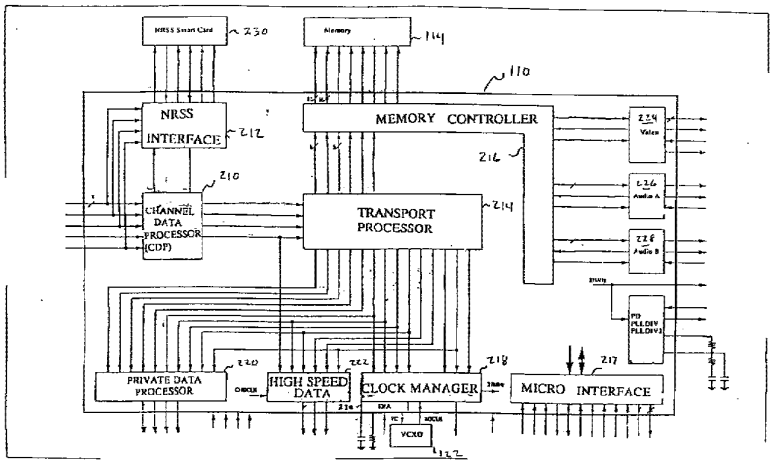


FIGURE 2A

EP 0 798 932 A2

Description

FIELD OF THE INVENTION

The present invention relates generally to data reception, processing and transmission of data encoded according to MPEG-2 standards and, more particularly, the present invention relates to the operation and interfacing of a transport decoder which handles an MPEG-2 format datastream.

BACKGROUND OF THE INVENTION

High Definition Television (HDTV) continues to make progress in its attempts to replace conventional television. Paving the way for this progress are various companies and associations working on standards to provide for a global market for HDTV.

One such group of companies is known as the "Digital HDTV Grand Alliance" including members such as AT&T, David Sarnoff Research Center, Massachusetts Institute of Technology and others. A comprehensive overview of the strides made by this group are presented in an article by Robert Hopkins entitled "Digital Terrestrial HDTV for North America: The Grand Alliance HDTV System" published in the IEEE Transactions on Consumer Electronics (Summer 1994). This article is herein incorporated by reference for all of its teachings regarding the background and basics of HDTV systems including the use of Program and Transport Packet Streams.

One standard that has been adopted by the Grand Alliance is the MPEG-2 standard for encoding video and audio information, developed by the Moving Pictures Expert Group (MPEG), a committee within the International Standards Organization (ISO). Accepted standards are periodically published such as, for example, the Video Section of Information Technology - Generic Coding of Moving Pictures and Associated Audio ISO/IEC 13818-2 (1995) (hereinafter "Video Section") and the Systems Section of Information Technology - Generic Coding of Moving Pictures and Associated Audio ISO/IEC 13818-1 (November 1994) (hereinafter "Systems Section") both of which are incorporated herein by reference for their teachings regarding established standards and formats.

The syntax for the MPEG-2 standard defines several layers of data records which are used to convey both audio and video data. To transmit information, a digital data stream, representing, for example, multiple video sequences, is divided into several smaller units and each of these units is encapsulated into a respective packetized elementary stream (PES) packet. For transmission, each PES packet is divided, in turn, among a plurality of fixed-length transport packets. Each transport packet contains data relating to only one PES packet. The transport packet also includes a header which holds control information, sometimes including an adaptation field, to be used in decoding the transport packet. However, PES data may include multiple elementary streams.

As set forth in the Systems Section, synchronization among multiple elementary streams is accomplished by periodically providing time stamps in the data stream. For instance, the decoding of N elementary streams is synchronized by adjusting the decoding of the streams to a common master time base rather than by adjusting the decoding of one stream to match that of another. The master time base may be the clock signal of one of the N decoders, or the clock signal of the data source, or it may be an externally supplied clock signal.

To aid in this synchronization process, included in the transport header of the MPEG-2 standard is a "discontinuity indicator". A syntax representation is provided in the Systems Section showing where in the transport header an exemplary "discontinuity indicator" resides. This syntax is re-presented below in Table I.

Table I -- Transport Stream adaptation field

Syntax	No.of Bits	Mnemonic
adaptation_field() {		
adaptation_field length	8	imsbf
if(adaptation_field_length > 0) {		
discontinuity_indicator	1	bslbf
random_access_indicator	1	bslbf
elementary_stream_priority_indicator	1	bslbf
PCR_flag	1	bslbf
OPCR_flag	1	bslbf
splicing_point_flag	1	bslbf
transport_private_data_flag	1	bslbf
adaptation_field_extension_flag	1	bslbf
if(PCR_flag == '1') {		
program_clock_reference_base	33	uimsbf
reserved	6	bslbf
program_clock_reference_extension	9	uimsbf
}		
if(OPCR_flag == '1') {		
original_program_clock_reference_base	33	uimsbf
reserved	6	bslbf
original_program_clock_reference_extension	9	uimsbf
}		
if(splicing_point_flag == '1') {		
splice_countdown	8	tcimsbf
}		
if(transport_private_data_flag == '1') {		
transport_private_data_length	8	uimsbf
for (i=0; i<transport_private_data_length;i++){		
private_data_byte	8	bslbf
}		
}		

Table I (continued)

Syntax	No. of Bits	Mnemonic
if(adaptation_field_extension_flag == '1'){		
adaptation_field_extension_length	8	uimsbf
ltw_flag	1	bslbf
piecewise_rate_flag	1	bslbf
seamless_splice_flag	1	bslbf
reserved	5	bslbf
if(ltw_flag == '1'){		
ltw_valid_flag	1	bslbf
ltw_offset	15	uimsbf
}		
if(piecewise_rate_flag == '1'){		
reserved	2	bslbf
piecewise_rate	22	uimsbf
}		
if(seamless_splice_flag == '1'){		
splice_type	4	bslbf
DTS_next_au[32..30]	3	bslbf
marker_bit	1	bslbf
DTS_next_au[29..15]	15	bslbf
marker_bit	1	bslbf
DTS_next_au[14..0]	15	bslbf
marker_bit	1	bslbf
}		
for (i=0;i<N;i++){		
reserved	8	bslbf
}		
}		
for (i=0);i<N;i++){		
stuffing_byte	8	bslbf
}		
}		
}		

As described in the Systems Section, the discontinuity indicator resides within an adaptation field within the Transport Header. The discontinuity indicator is a 1 bit field which, when set to a logical '1', indicates that the discontinuity state is true for the current Transport Stream packet. When the discontinuity indicator is set to a logical '0' or is not present, the discontinuity state is false. The discontinuity indicator is used to indicate two types of discontinuities, system time-base discontinuities and continuity counter discontinuities.

A system time-base discontinuity is indicated by the use of the discontinuity indicator in Transport Stream program identification (PID) packets designated as a Program Clock Reference (PCR) PID (refer to Systems Section 2.4.4.9 on page 50). Subsequently, a PCR value (i.e., reference to a master clock signal) is received.

The discontinuity indicator bit is set, for example, when program material which was recorded with one time base is inserted into a data stream that was recorded with another time base. The difference in time base appears as an unexpected PCR value.

When the discontinuity state is true for a Transport Stream packet of a PID designated as a PCR PID, the next PCR in a Transport Stream packet with that same PID represents a sample of a new system time clock for the associated program. The system time-base discontinuity point is defined to be the instant in time when the first byte of a packet containing a PCR of a new system time-base arrives at the input of the System Target Decoder (STD).

It should be noted that an STD is well known in the art and, briefly, represents a hypothetical decoder which is conceptually connected to an output of a transport encoder. It is required that a datastream that conforms to the MPEG-2 standard shall not cause dataflow problems for the STD (including buffer overflow). Additional discussion of the STD is

provided in the Systems Section.

Continuing, the discontinuity indicator bit may also be set to '1' in Transport Stream packets of the same PCR PID prior to the packet which contains the new system time-base PCR. In this case, once the discontinuity indicator has been set to '1', it continues to be set to '1' in all Transport Stream packets of the same PCR PID up to and including the Transport Stream packet which contains the first PCR of the new system time-base. After the occurrence of a system time-base discontinuity, no fewer than two PCRs for the new system time-base will be received before another system time-base discontinuity can occur. Further, except when a trick mode status is true, data from no more than two system time-bases shall be present in the STD buffers for one program at any time.

A potential problem occurs when a time-base discontinuity requires that multiple System Time Constant (STC) values (e.g., in a Clock Manager of a Transport Decoder and in external video and audio decoders) be updated, particularly when a substantial latency may exist between data that is presently being decoded by external decoders and data that is being received and stored in the buffers of a transport decoder. That is, typically a transport decoder is coupled to external video and audio decoders each of which maintain an STC or equivalent thereof. Thus, when a change in system time-base occurs, adjusting all of the STC's at the same time may create system synchronization problems.

With this in mind, an important concern, addressed by the present invention, is how a transport decoder can efficiently handle the occurrence of a system time-base discontinuity between buffers and decoders.

SUMMARY OF THE INVENTION

The present invention involves a system designed for receiving a transport data stream which includes transport packets having headers and payloads, the headers include a discontinuity indicator. The system including at least one buffer for temporarily storing the received transport data stream and at least one external decoder for decoding data output from the buffer for presentation purposes. The present invention is designed to handle the receipt of a discontinuity indicator by, upon receiving a discontinuity indicator, parsing the transport packet header and storing the transport packet payloads in memory. Upon receiving a next program counter reference value, a counter is loaded with the received program counter reference value. The data stored is then searched for a time stamp and, when one is found, the time stamp is retrieved. Finally, a timer interrupt is set for a point in time based on the retrieved time stamp, and, when the timer interrupt occurs, the program counter reference value is sent to the decoder, providing the new time stamp value insufficient time for the decoder to process data having the new time base without interfering with the processing of data having the previous time base.

BRIEF DESCRIPTION OF THE DRAWINGS

The invention is best understood from the following detailed description when read in connection with the accompanying drawing, in which:

Figure 1A shows a high-level functional block diagram of an exemplary digital transport decoder and its various interfaces.

Figure 1B shows a high-level flowchart illustrating the steps followed by one aspect of the present invention using the exemplary implementation shown in Figure 1A.

Figure 2A shows a high-level functional block diagram of an exemplary implementation of a the transport decoder shown in Figure 1A.

Figure 2B shows a high-level data/control flow block diagram of an exemplary implementation of the transport decoder shown in Figure 1A.

Figure 3 shows a functional block diagram of the memory controller used in the transport decoder shown in Figure 2.

Figure 4 shows an exemplary flowchart illustrating how the memory controller handles Pointers FIFO Entries.

Figure 5A shows an exemplary flowchart illustrating steps executed during the video decoder interface aspect of the present invention.

Figure 5B shows an exemplary flowchart illustrating steps executed during the audio decoder interface aspect of the present invention.

DETAILED DESCRIPTION OF THE INVENTION

Overview

As mentioned above, one important aspect of the present invention is how a transport decoder can efficiently handle a discontinuity indicator. Another aspect of the present invention is interfacing a transport decoder to a video decoder which can only handle elementary stream data. Still another aspect of the present invention involves interfacing a transport decoder to a bit rate-constrained audio decoder. To encompass the above listed aspects, the present invention includes a specially-designed MPEG-2 Transport Decoder.

To provide some context for the discontinuity aspect of the present invention, a real-world example of a time-base discontinuity would occur if a program were being broadcast, using MPEG-2 formats, via satellite to many receiving stations across the country. But, during the program, each station inserted its own respective commercials. In this situation, a discontinuity indicator would be set to indicate a gap or break in the time-base of the program in order to allow for the insertion of commercials, each having its own local time base. The program transmission from the satellite, however, would continue while the individual stations inserted their respective programs. Accordingly, the internal timing of the transport decoder for receiving the program broadcast from the local station with the inserted commercials would desirably track the appropriate time-bases when a discontinuity indicator is encountered.

Turning to the figures, Figure 1A shows a high-level functional block diagram of an exemplary digital transport decoder 110 and its various interfaces. As shown in Figure 1A, transport decoder 110 includes connections to a physical layer channel interface (channel interface) 112, buffer memory 114, a host microprocessor 116, external video and audio decoders 118 and 120, and clock signal circuitry (e.g., VCXO) 122. In the exemplary embodiment of the present invention, the transport decoder 110 supports one video interface and up to two audio interfaces (e.g., audio A and audio B).

Host microprocessor 116 is also coupled, external to transport decoder 110, to the video and audio decoders 118 and 120 by way of microbus 124. This connection may use parallel or serial data paths and depends on the requirements of the individual external decoders selected for use. In any event, as will be appreciated by those of ordinary skill in the art, this method of coupling the host microprocessor 116 to external decoders is well known.

Generally, the transport decoder 110 and each of the external video and audio decoders 118 and 120 include a System Time Constant (STC) value (not shown) maintained in a register or counter for the purpose of synchronizing the decoding and display of the received data stream.

As mentioned, a problem can occur when a time-base discontinuity occurs within the data stream, and the discontinuity requires STC values to be updated both in the transport decoder 110 and in the video and audio decoders 118 and 120. This is particularly a problem when a substantial latency may exist between data presently being decoded by the external decoders and data being received and stored in the transport processor 110.

In the present invention, microprocessor 116 has an important function in controlling the synchronization of the datastream and the operation of the video and audio decoders 118 and 120. In the present invention, microprocessor 116 has direct, memory-mapped access to various on-chip registers within the transport decoder 110. Microprocessor 116 also has access to buffer areas in the external memory 114 (through read, write and watermark pointers) and to on-chip 32-byte Read and Write FIFOs (shown in Figure 3). Microprocessor 116 is coupled to several internal functional blocks by an internal microbus 223 (shown in Figure 2B) and allows for various event and error conditions to be signaled via maskable interrupts.

Generally, as illustrated in the flowchart of Figure 1B, one aspect of the present invention, by way of the specially-designed transport decoder 110, performs the following steps after receiving of a discontinuity indicator, step 150:

1) because a discontinuity indicator is not necessarily in the same data packet as a next Program Clock Reference (PCR) value, when the next PCR value is received, a counter within Clock Management circuitry (shown in Figure 2A) contained within transport decoder 110 is loaded with the next PCR value, step 152;

3) subsequent packets, parsed and stored in memory, are searched for a time stamp and, when one is found, the time stamp value is retrieved, step 154;

4) a timer interrupt is set by the microprocessor 116 for a point in time that is just less than one frame time before the retrieved time stamp value, step 156; and

5) when this timer interrupt occurs, the current STC value is sent to the external video and audio decoders 118 and 120, step 158, thereby giving the external decoders sufficient time to process the data before the time stamp value in the datastream is compared to the counter value maintained in the video or audio decoders 118 or 120.

The implementation of these high level steps is described in detail below.

Figures 2A and 2B show high-level functional block diagrams of an exemplary implementation of the transport decoder 110 shown in Figure 1A. With reference to Figures 2A and 2B, the overall data flow in the transport decoder 110 is described.

As shown in the figures, the channel interface 112 of Figure 1A leads to a channel data processor (CDP) 210. Channel interface 112 also provides data to an NRSS interface 212 for purposes of decryption by the NRSS Smart Card 230. The EIA Standard for Conditional Access, Version 2.6, NRSS Committee (4/95) is hereby incorporated by reference for its teaching regarding the NRSS Smart Card.

The CDP 210 is provided to preprocess (e.g., detect, synchronize and validate) the packets in the received datastream. The CDP 210, after performing its processing, forwards its output data to transport processor (TPROC) 214 for further processing. TPROC 214 performs various processing, such as parsing, in accordance with a PID Table described below in detail.

The TPROC 214 interacts with several of the functional blocks within transport decoder 110 including a Memory Controller 216, a Clock Manager 218, a Private Data Processor 220, and a High Speed Data Port 222. As mentioned, the microprocessor 116 has direct access to various on-chip registers. This access is gained by way of a Micro Interface 217 and the internal microbus 223 (shown in Figure 2B). Memory Controller 216, in addition to being coupled to TPROC 214, is also coupled to external memory 114 and video and audio decoders 224, 226 and 228.

The details of the processing of the digital datastream, with respect to the individual functional blocks, is now described.

Channel Data Processor (CDP) 210

It is noted that the datastream from the channel interface 112 is passed to the CDP 210, after possible buffering (not shown).

The CDP 210 performs various preprocessing steps on the received datastream. In particular, CDP 210 determines the byte and frame boundaries in the datastream and converts the received datastream into an 8-bit parallel format. In the exemplary embodiment of the present invention, operations occurring in the transport decoder 110 subsequent to the CDP 210 are performed using a byte format.

To determine byte and frame boundaries, the CDP 210 synchronizes to a framing pattern contained in the header of a Transport Packet. In particular, this synchronization is achieved by searching for the MPEG-2 Transport Stream sync_byte (e.g., 47H) and verifying its occurrence at the start of a programmable number of consecutive transport packets.

In doing so, the CDP 210 checks for the sync_byte, "locks" on to the datastream and continues validating packets. When the sync_byte search is in progress, a Search_State bit is set in a CDP Status Register (not shown). When the sync_byte has been found and verification of successive occurrences is in progress, a Sync_State bit is set. Finally, a Lock_State bit is set when a predetermined search criterion is met. In the exemplary embodiment of the present invention, the search criterion is three consecutive validated packets. Also, a Sync_Hysterisis value, indicating the number of packets with corrupted sync_bytes that the CDP 210 is programmed to allow to pass through before declaring loss of Lock_State, is established. When this hysterisis value is exceeded, a loss of lock is declared. The CDP 210 then issues a Lock_Lost interrupt and transitions from the Lock_State back to the Search_State.

Additionally, if the channel interface 112 (shown in Figure 1A) provides a CHPSTRT signal, the transport decoder 110 may be configured to use this signal as a means to synchronize to the datastream by setting a Framing_Mode bit in a CDP Framing and Sync Control Register to logical "1". In this case, synchronization is achieved within one transport packet time after the CHPSTRT is asserted.

Alternatively, if an Ignore_Sync bit, in the CDP Framing and Sync Control Register, is set to logical "0", the CDP 210 verifies sync_byte integrity at the expected (first) byte location relative to the CHPSTRT signal transition and declares a lock condition if verification falls. If Ignore_Sync is set to 1, the CDP 210 does not verify sync integrity.

After achieving the Lock_State, the output of the CDP 210 can be passed to and processed by the TPROC 214, which further processes (e.g., filters, parses, and formats) the data now in byte format.

Transport Processor (TPROC) 214

Signals produced by the TPROC 214 control the flow of data to the Memory Controller 216, the Micro Interface 217, the Clock Manager 218, the Private Data Processor 220 and the High Speed Data Port 222.

In the exemplary embodiment of the present invention, the TPROC 214 processes the data stream under control of a PID Table. As described in the Systems Section, a PID is a 13-bit field in a Transport Stream Header, indicating the type of data stored in the packet payload. Some PID values are assigned and some are reserved. Additional details of the PID Table are described in Section 3.3 of the Systems Section of the above-referenced MPEG-2 specification.

Regarding the PID Table, the TPROC 214 is capable of concurrently processing up to 32 user-selectable PIDs. The PIDs which the transport decoder 110 is to process are specified in the PID Table. Bit fields in the PID Table specify

various processing options for the data.

In operation, the PID is extracted by the TPROC 214 from the header of an incoming transport packet and simultaneously compared to all of the entries in the PID Table. If a match is found, the data in the packet is processed according to the options set in the matching PID Table entry. If no PID Table entry matches, the packet is discarded.

Each location in the PID Table for which the Payload Format bits are not set to Discard (000), is associated with a channel number. In conjunction with the TPROC 214, the Memory Controller 216 is responsible for transferring data to a specific buffer area (channel) in RAM. Operation of the Memory Controller 216 is discussed in detail below with reference to Figures 2, 3 and 4.

In the exemplary embodiment of the present invention, the PID Table, maintained in TRPROC 214, is set up under software control by the microprocessor 116 via the internal microbus 223 during an initialization process.

Regarding PID Table entry format, a PID Table entry includes, among other information, 1) a 13-bit PID and 2) a 1-bit PCR PID field which, when set, indicates that this PID carries the PCR for the program being decoded. An exemplary PID Table entry also includes a 1 bit PES HDR bit which, when set, enables the RAM buffer address corresponding to a "Stream ID" in the PES header to be stored to a Pointers FIFO. An interrupt is issued by the Memory Controller 216 when it writes data to the Pointers FIFO. This interrupt is referred to as a DMA_MARK signal which serves to alert the microprocessor 116 of the delivery of the Stream_ID byte. In the exemplary embodiment of the present invention, Channel Numbers are assigned to PID Table entries in numerical order based on their location in the table, for example, the first entry is assigned DMA Channel 0, the next one DMA Channel 1 and so on.

The TPROC 214 indicates that transport packets with PCR PID should be processed by the Clock Manager 218 by placing a PCR_PID signal in the PID Table entry. The Clock Manager 218 then monitors the adaptation fields of such packets for a PCR_flag and extracts the PCR.

After separating the packets according to their PIDs, the TPROC 214 removes the headers of the transport packets and, in conjunction with the Memory Controller 216, stores packet payloads (e.g., portions of PES packets) for the respective PID's in sequential memory locations of the designated channel in the external memory 114. The data in the memory 114 represents PES packets.

The TPROC 214 also handles write operations to the external memory from the host microprocessor 116.

MEMORY CONTROLLER 216

The Memory Controller 216 stores the parsed payload of Transport Packets in external memory 114 and provides video, audio, PSI and other data (e.g., private data) upon demand to the Video Processor 224, up to two Audio Processors 226 and 228 and to the host microprocessor 116 (or microcontroller).

Figure 3 shows an exemplary functional block diagram of a Memory Controller suitable for use in the transport decoder 110 shown in Figures 2A and 2b. As shown in Figure 3, the Memory Controller 216 receives data from the TPROC 214 via a Transport In FIFO 310. Memory Controller 216 receives/transmits data from the microprocessor 116 via Micro Interface 217 including Micro-in FIFO 314 and Micro-out FIFO 316. It is noted that the Micro Interface 217, shown external to Memory Controller 216 in Figures 2A and 2B, is shown in Figure 3 for the sake of clearly illustrating its interrelationship with FIFOs 314 and 316.

The Memory Controller 216 also includes a Flow Control Unit 318 containing On Chip Memory, a DMA Controller 320, Datapath logic 322, and FIFOs 324, 326 and 328 for delivering data to the video and audio decoders 224, 226 and 228, respectively.

The Memory Controller 216 writes data into external memory from two sources: one is the TPROC 214 (through the Transport In FIFO 310) and the other is the microprocessor 116 (through the Micro In FIFO 314).

The Memory Controller 216 reads data from external memory and supplies it to the following four destinations 1) Video Interface (through the Video FIFO 324), 2) Audio A Interface (through the Audio A FIFO 326), 3) Audio B Interface (through the Audio B FIFO 328), and 4) microcontroller 116 (through the Micro Out FIFO 316).

The external memory space is divided into separate, non-overlapping buffer areas, one for each DMA Channel in use. Up to 32 channels can be used. A channel number is associated with data written into the two source FIFOs or read out of the four destination FIFOs.

Data provided by the TPROC 214 and written into the Transport In FIFO 310 carries a channel number tag. The channel number for each of the other FIFOs is defined by the microcontroller 116 in a DMA Interface Register. A 128-bit DMA Channel Configuration Register (Channel Register) is set up by the microprocessor 116 for each channel in use. The Channel Register contains configuration information such as the buffer space defined by Begin and End Addresses, whether the buffer is configured as a FIFO (read and write pointers wrap around, sometimes referred to as a circular queue) or as a Queue (read and write pointers do not wrap around), etc.

If the Channel Register allows, the DMA Controller 320 can obtain a new buffer when the one it is currently using becomes full. This feature is normally used only for buffers which store data that is destined for the host microprocessor 116, such as PSI sections. The Begin and End Addresses of these new buffers are stored in a New Buffer FIFO. Only the microprocessor 116 can write new buffers to this FIFO. Used buffer pointers and other types of pointers are written

by the DMA Controller 320 to the Pointers FIFO. In the exemplary embodiment of the present invention, only the microprocessor 116 can read the Pointers FIFO. These pointers are used by the microprocessor 116 to manage reading of buffer data and to return used buffers back to the New Buffer FIFO. The Pointers FIFO Entries Flowchart, shown in Figure 4, shows the entries to the Pointers FIFO for all possible signal conditions, Channel Register configurations and buffer conditions.

Continuing with Figure 3, the DMA Controller 320 and Datapath logic 322 provide all interface signals to the external memory. The Flow Control Unit 318 directs the DMA Controller 320 to service a specific interface FIFO (one of six) by providing the appropriate control signals to it. These include:

1. DMA Controller Start

This signal commands the DMA Controller 320 to start servicing a specific FIFO using the commands such as DMA Controller Operation, DMA Input FIFO Select and DMA Output FIFO Select. At the same time, Flow Control Unit 318 reads out the 128-bit Channel Register and makes it available on the data bus of the Channel Register (internal memory). The DMA Controller 320 reads this 128-bit word into its internal register and the Flow Control Unit 318 is then free to do other read/write operations as required on the internal memory.

2. DMA Controller Operation (1 bit)

This bit signals a Write or Read operation to be performed by the DMA Controller on the interface FIFOs.

3. DMA Input FIFO Select (1 bit)

Signals which of the two input FIFOs 310 or 314 to service.

4. DMA Output FIFO Select (2 bits)

Signals which of the four output FIFOs 316m 324, 326 or 328 to service.

5. Channel Data In Request/Channel Data/Channel Number/DMA Controller Stop:

Usually, the DMA Controller 320 continues to service a FIFO until it cannot service the FIFO any longer due to some condition such as a full interface FIFO, or no data in external memory, etc. When this type of condition occurs, the DMA Controller 320 asserts the Channel Data In Request to the Flow Control Unit 318 and provides the Channel Register data and Channel Number data to the Flow Control Unit 318. The Flow Control Unit 318 updates the Channel Register memory with the Channel Data (updated information includes, for example, the Write and Read Pointers) and directs the DMA Controller 320 to service the next FIFO in the loop, unless a specific FIFO needs emergency action (because, for example, it may be approaching a full or empty status).

Occasionally, the Flow Control Unit 318 may need to issue a DMA Controller Stop command to the DMA Controller 320 in response to a need to service another FIFO such as the Transport In FIFO which may have a higher priority. This command directs the DMA Controller 320 to stop servicing the FIFO it is currently servicing and write the Channel data to the Flow Control Unit 318.

The Channel Register is updated by the DMA Controller 320 with a Read Pointer, a Write Pointer and other information at the point at which it stops servicing a FIFO. Since the Flow Control Unit 318 writes this to the internal Channel Register memory, the DMA Controller 320 can resume service to the interrupted FIFO from the point at which it was interrupted.

With reference to the DMA Controller's participation in the overall operation of the discontinuity state aspect of the present invention, when a next PES packet containing a Stream ID byte is written to the Memory Controller 216 by TPROC 214, the Flow Control Unit 318 causes the DMA_MARK signal to go high. The transition of the DMA_MARK signal triggers an interrupt at microprocessor 116 and the address of the PES packet header is stored in an input pointer FIFO to which the microprocessor has access. Since this PES header may have a time-stamp, the microprocessor 116, responsive to the flag raised by the discontinuity indicator, searches the stored PES headers for a time stamp and, when it finds one, retrieves the time stamp value. Based on the retrieved timestamp value, microprocessor 116 sets a timer interrupt.

Referring back to Figure 3, in the exemplary embodiment of the present invention, internal memory of the Flow Control Unit 318 includes four blocks of 64x32 bits, each of which are assigned for Channel Registers, New Buffer FIFO and Pointers FIFO.

In the exemplary embodiment of the invention, each Channel Register has 128 bits and there may be a maximum of 32 channels, so up to 128x32 bits can be assigned to Channel Registers, in multiples of 128 bits.

Words of 32-bits are written to and read from the New Buffer and Pointers FIFOs. At least 128x32 bits of memory (i.e., 128 words) are available for the New Buffer and Pointer FIFOs. Additionally, if less than 32 channels are in use, unused capacity from the Channel Register memory can be allocated to the New Buffer and Pointer FIFOs.

Following are examples of memory allocation for 32 DMA channels and 16 DMA channels.

1) Memory allocation for 32 used channels

New Buffer FIFO=64x32 bits
Pointer FIFO=64x32 bits
Channel Registers=128x32 bits

2) Memory allocation for 16 used channels

New Buffer FIFO=64x32 bits
Pointer FIFO=128x32 bits
Channel Registers=128x16 bits

CLOCK MANAGER 218

The Clock Manager 218, shown in Figures 2A and 2B, of the transport decoder 110 synchronizes a local 27 MHz signal (generated by clock signal circuitry 122) using successive PCR values. The value of the counter in the Clock Manager 218 is referred to as the System Time Constant (STC) value. The microprocessor 116, in the exemplary embodiment of the present invention, has access to the STC value in the counter via the internal microbus 223.

The counter in the Clock Manager 218 is periodically synchronized to the PCRs carried in the transport data stream. When the first PCR in a data stream is encountered, the PCR is "jam loaded" in the counter to become the initial STC value. As subsequent PCR values are received, they are compared to the STC value to adjust the frequency of the clock signal circuitry 122 (e.g., VCXO). In the present invention, this is accomplished using a Digital-to-Analog converter (DAC) described in detail below. Once phase error adjustments are made, generally, the new PCR value is loaded as the STC value.

Normally PCR values occur once per frame. Under the MPEG-2 standard, they must occur at least in every third frame (one-tenth of a second).

The external video and audio decoders 118 and 120 are typically synchronized to the transport decoder 110 by updating the STC value in the external decoders based on the STC value maintained in the Clock Manager 218. The video and audio decoders 118 and 120 use the stored value to determine when data is to be decoded/displayed by comparing it to timestamps in the received PES datastream. It is noted that, in the exemplary embodiment of the present invention, the transport decoder 110, video decoder 118 and audio decoder 120 all receive the same 27 MHz clock signal generated by the Clock Manager 218.

The transport decoder 110 contains hardware support for locally generating a clock signal that is locked to the system clock signal for a selected program. As mentioned above, PCR values received on the specified PCR_PID are made available on a serial interface for use by other devices.

The Clock Manager 218 includes a System Time Counter (STC) (not shown), a System Time Clock Register (not shown), a Program Clock Reference Register(not shown), a Current STC Register (not shown), an Alarm Clock Interrupt Register (not shown) and logic (not shown) to latch PCRs extracted from the headers that are associated with a selected PID. It also contains the serial PCR interface, and a DAC Register for control of an external system clock signal control loop.

In the exemplary embodiment of the present invention, the STC is a 42-bit counter having a 9-bit base and 33-bit extension.

The STC base divides the nominal 27 MHz System Clock input signal by 300. The resulting 90 KHz signal drives the STC extension. The STC can be read or written by software executing on the microprocessor 116 at any time. Optionally, the STC can be automatically loaded with a received PCR after a discontinuity indicator has been encountered.

As mentioned above, PCRs may be automatically extracted from the incoming bit stream by setting the PCR_PID bit in a PID Table entry. Handling of the received PCRs depends upon options set by software and whether a discontinuity state exists for the current transport packet when the PCR is received.

A discontinuity state is recognized when a transport packet which contains the PCR_PID is received with a logic 1 in the discontinuity_indicator bit in the Transport header. The discontinuity state exists until either a transport packet containing the PCR_PID is received with a logic 0 in the discontinuity_indicator, or until the next PCR is found in a PCR_PID packet.

If enabled, the Clock Manager 218 automatically loads the received PCR into the STC when a PCR is received

while a discontinuity state exists. At this instant, an interrupt is issued to the host microprocessor 116.

The Clock Manager 218 contains an Initialize bit which, when set to logic 1 by software, causes the next PCR received to be loaded into the STC regardless of whether a discontinuity state exists or not. The Initialize bit is subsequently cleared automatically by the Clock Manager 218. The Initialize bit is set only when the STC does not contain a valid PCR (i.e., when a new PID has been selected.)

If no discontinuity state exists or if the automatic STC loading function is disabled, a received PCR is stored in the PCR register and (optionally) an interrupt is issued to the host microprocessor 116. At the same time that the PCR is captured, the current state of the STC is stored in the STC Register. The stored PCR and STC values are used to compute an error signal to control the external system clock signal (XOCLK).

When a Transport Packet is received which has a PID equal to the PCR_PID and which contains a PCR, the PCR is latched in the PCR Register. The current contents of the STC Counter are also latched in the STC Register at that time, and an interrupt is issued to the host microprocessor 116. Latching both values eliminates any uncertainty due to varying latency in the interrupt service routine. In the exemplary embodiment of the present invention, both registers consist of a 9 bit base and 33-bit extension. The base part is modulo 300, whereas the extension is binary. Before the register contents can be used in any calculations, they are converted to a single binary count by multiplying the extension by 300 and adding the base part to the result.

The Current STC Register is a read-only register which allows the microcontroller 116 to obtain the current contents of the STC at any time. The microcontroller 116 can initialize the STC of the Video Decoder or Audio Decoder with this value by way of the external microbus 124.

As described above, transport decoder 110 contains a 10-bit digital-to-analog converter (DAC) 230 for use in the clock adjustment control loop within the Clock Manager 218. The DAC 230 is write-only, i.e., software keeps a copy of the last value loaded into the DAC in memory. In operation, the DAC 230 is used to generate the control voltage for the external resonant crystal voltage controlled oscillator, VCXO 122, which is source of the system clock frequency (27 MHz nominal). By comparing the latched PCR and STC values, software can determine whether the local clock signal requires adjustment. For example, software may take the difference between the PCR and STC and add this error term to the current contents of the input register of DAC 230.

To review, the discontinuity aspect of the present invention includes:

1) The Channel Data Processor (CDP) 210 checks the synchronization byte in received transport packets to validate the packets. The Transport Processor (TPROC) 214 parses the transport header and passes the packet payloads to the proper location in the memory. Data is directed to different FIFO buffers (channels) in the memory based on the PID values in the transport headers.

2) Because a received discontinuity indicator is not necessarily in the same packet as the next PCR value, when the next PCR value is received, the TPROC 214 "jam loads" the counter in the Clock Manager 218.

3) When the next PES packet containing a Stream ID byte is written to the Memory Controller 216 by TPROC 214, the DMA_MARK signal goes high. This causes the Memory Controller 216 to interrupt the microprocessor 116 and store the address of the PES packet header in an input pointer FIFO. This PES header may have a time-stamp (either decode timestamp or presentation timestamp). The microprocessor 116, responsive to the interrupt from the Memory Controller 216, searches the PES headers for a time stamp and, when it finds one, retrieves the time stamp value.

4) The microprocessor 116 then sets a timer interrupt for a time that is just less than one frame time before the timestamp value.

5) When this timer interrupt occurs, the microprocessor 116 sends the new PCR value to the video and audio decoders by way of the external microbus 124. This gives the decoders sufficient time to process the data before its time-stamp is compared to the counter value in the video or audio decoder, without interfering with the processing of data that is reverenced to the prior PCR value.

VIDEO/AUDIO DECODERS 224, 226, 228

Once the transport packets are received, processed and stored, as mentioned above, Memory Controller 216 reads the PES packets out of memory 114 and forwards them to the external video and audio decoders. Before the PES packets reach the external decoders 118 and 120, however, they are processed by internal video and audio processors 224, 226 and 228.

An important role of these internal processors is to ensure that the stream of data leaving the transport decoder is compatible with the external decoders. That is, different external decoders have different data requirements.

For example, some currently available external video decoders can accept a stream of PES packets, including header and payload, essentially in the same form in which it is read out of memory. On the other hand, some external decoders can only accept for processing an elementary stream (i.e., PES packet payloads but no headers). Additionally, the latter type of decoder may require, in addition to the datastream, associated control information such as timestamps and byte counts. One such video decoder is the CL9100 made by C-Cube, the specification of which -- CL9100 Multimode Video Decoder User's Manual (10/94) -- is hereby incorporated by reference. Accordingly, video processor 224 is configurable, by way of microprocessor 116, to format the outgoing datastream to be compatible with the requirements of the selected external decoder.

With respect to handling the interface to an external video decoder such as a C-Cube CL9100, video processor 224 processes the stream of PES packets in order to strip off the PES header and extract, from the PES header, timestamp information. In addition, video processor 224 maintains a byte count, using a counter (not shown), for the PES packets forwarded to the external decoder. Thus, video processor 224 holds timestamp information, maintains a byte count and forwards the PES packet payloads on to the external video decoder. In order to get the timestamp and byte count information to the external decoder, video processor 224 sends an interrupt to microprocessor 116 which, when it responds, reads the timestamp and byte count information from the video processor 224 via the internal microbus 223 and delivers this control information to the external video decoder via the external microbus 124.

Sometimes, however, before microprocessor 116 is able to respond to the interrupt from video processor 224, another PES header with timestamp information arrives. This may occur, for example, if the microprocessor 116 was involved with a higher priority task for an extended period of time. In this case, video processor 224 stores the timestamp information along with a status. In particular, video processor 224 maintains a status register indicating whether microprocessor has "missed" any timestamps. When, in the exemplary embodiment of the invention, microprocessor 116 finally begins to read the timestamp information, if it has missed one or more timestamps as indicated in by the status, it only retrieves the most recent timestamp.

Additionally, since timestamp information and byte count information are to be provided to the external decoder, when the microprocessor 116 misses a timestamp, video processor 224 maintains a cumulative byte count for however many timestamps may arrive before microprocessor 116 reads the information. Thus, microprocessor 116 reads the most recent timestamp and cumulative byte count and delivers this information to the external video decoder. Subsequently, the status register and byte count may be reset.

As shown in the flowchart of Figure 5A, the headers of received data are stripped and parsed (step 510), timestamp information is stored and a byte count is maintained (step 512), an interrupt is issued by video processor 224 to microprocessor 116 alerting it that timestamp information is available for reading (step 514), if additional timestamp information is received before microprocessor reads the previous timestamp information, the new timestamp information is stored, the byte count continues to accumulate (step 516) and the video processor 224 issues another interrupt to the microprocessor 116, otherwise microprocessor reads the status register, the timestamp information and the byte count information and delivers it to the external video decoder (step 518).

Another example of accommodating particular external decoders involves the use of bit rate-constrained audio decoders such as the Zoran ZR38500 AC3 audio decoder, the Specification (10/94) for which is hereby incorporated by reference. Generally, most audio decoders present a data request indicating they can accept audio data and then the transport decoder 110, via the audio processors 226/228, supplies audio data until the external audio processor negates the requests. In other words, there exists a handshake mechanism.

Bit-rate constrained decoders, however, such as the Zoran AC3 audio decoder, require that audio AC3 data be delivered as audio data frames, each frame including a frame time and bit-rate in its header. Delivering data in excess of this predetermined bit rate can cause the internal buffers to overflow and, consequently, cause a loss of data. Furthermore, these bit-rate constrained decoders do not provide a Data-Ready signal to indicate that they are ready to accept additional audio data.

Thus, as with the video decoder 224, the audio processors 226 and 228, are configurable to accommodate the selected external audio decoder. It is noted that, in the exemplary embodiment of the present invention, audio processors 226 and 228, except for addresses which allow access by microprocessor 116, are essentially identical.

Accordingly, audio processors 226/228, when programmed to interface with this type of audio decoder, assumes that the decoder is always issuing a data request signal. However, as PES packet audio data is supplied to audio processors 226/228 from Memory Controller 216, the PES header of the received data is parsed by the audio processors. Contained in the parsed header is a frame size and bit rate for the audio data. From the frame size and bit rate, the audio processors 226/228 can determine the frame boundary (i.e., frame time). Having frame time information, audio processors 226/228 set a frame timer such that, when audio data is available, the timer is set, a frame of data is "burst" out to the external audio decoder then the audio processors 226/228 and then pause until the frame timer expires. When the timer expires, the cycle continues, one audio frame at a time, while additional frames are available for sending to the external audio decoder. Control information, such as timestamps, are retrieved by the microprocessor and delivered via the external microbus.

As shown in the flowchart of Figure 5B, PES packets are retrieved from memory (step 542), the headers of received

data are stripped and parsed (step 530), a frame time is determined (step 532), the timer begins running (step 534), audio processors 226/228 burst out a frame of data (step 536) and audio processors 226/228 pause until the timer expires (step 538), although, as shown, steps 542, 530 and 532 can be carried out during the pause. Then, the process can continue.

Although the invention is illustrated and described herein as embodied in a method and apparatus for updating the system time constant following a discontinuity in an MPEG-2 transport data stream, the invention is not intended to be limited to the details shown. Rather, various modifications may be made in the details within the scope and range of equivalents of the claims and without departing from the spirit of the invention.

Claims

1. A method for use in a transport decoder system designed for decoding a transport data stream which includes transport packets having headers and payloads, the headers including a discontinuity indicator, said system including at least one buffer for temporarily storing the received transport data stream and at least one decoder for decoding data output from the buffer for presentation purposes, each of said transport decoder and said decoder having separate counters, the method being a method for handling the receipt of a program counter reference (PCR) after a discontinuity state has been set comprising the steps of:

   parsing received transport packets and storing the data in memory;

   upon receiving a next PCR value, loading the transport decoder counter with the received program counter reference value;

   searching the data stored in memory which is received after the PCR value for a time stamp and retrieving the time stamp;

   setting a timer interrupt based on the retrieved time stamp value; and

   when the timer interrupt occurs, sending the PCR value to the counter of the decoder, thereby providing sufficient time for the decoder to process data.

2. The method of claim 1, wherein the timer interrupt is set for just less than one frame time before the time stamp value.

3. The method of claim 1, wherein the step of searching includes

   upon receipt of a transport packet which may contain a time stamp value, interrupting a host microcontroller to send the host microcontroller address information relating to the stored transport packet;

   subsequently, searching at the appropriate address location of the stored transport packet for a time stamp value.

4. The method of claim 1, further comprising the step of:

   before parsing, validating received transport packets by determining whether their boundaries correspond to predetermined packet boundaries and parsing only those packets which have been validated.

5. The method of claim 4, wherein the step of validating the received transport packet includes checking a synchronization byte in the received transport packet.

6. In a system designed for receiving a transport data stream which includes transport packets having headers and payloads, the headers including a discontinuity indicator, said system including at least one buffer for temporarily storing the received transport data stream and at least one decoder for decoding data output from the buffer for presentation purposes, each of said transport decoder and one decoder having separate counters, an apparatus for handling the receipt of a discontinuity indicator comprising the steps of:

   means for receiving and detecting a discontinuity indicator;

   means for parsing the transport packets and storing the data in memory;

means, upon receiving a next program counter reference value, for loading the transport decoder counter with the received program counter reference value;

means for searching the data stored in memory for a time stamp and, when one is found, for retrieving the time stamp;

means for setting a timer interrupt based on the time stamp value; and

means, when the timer interrupt occurs, for sending the program counter reference value to the counter of the one.

7. In a transport decoder system designed for decoding a transport datastream which includes transport packets having headers and payloads, the headers including a discontinuity indicator, said transport decoder coupled to at least one external decoder for decoding data output from the transport decoder, each of said transport decoder and said external decoders having separate counters for maintaining synchronization to the datastream, a method for handling the receipt of a discontinuity indicator comprising the steps of:

receiving the discontinuity indicator;

parsing the transport packet and storing the data in memory;

upon receiving a next program counter reference value, loading the transport decoder counter with the received program counter reference value;

searching the data stored in memory for a timestamp and, when one is found, retrieving the timestamp;

setting a timer interrupt for a point in time approximately one frame time before the time stamp value; and

when the timer interrupt occurs, sending the program counter reference value to the counter of the decoder, thereby providing sufficient time for the decoder to process data.

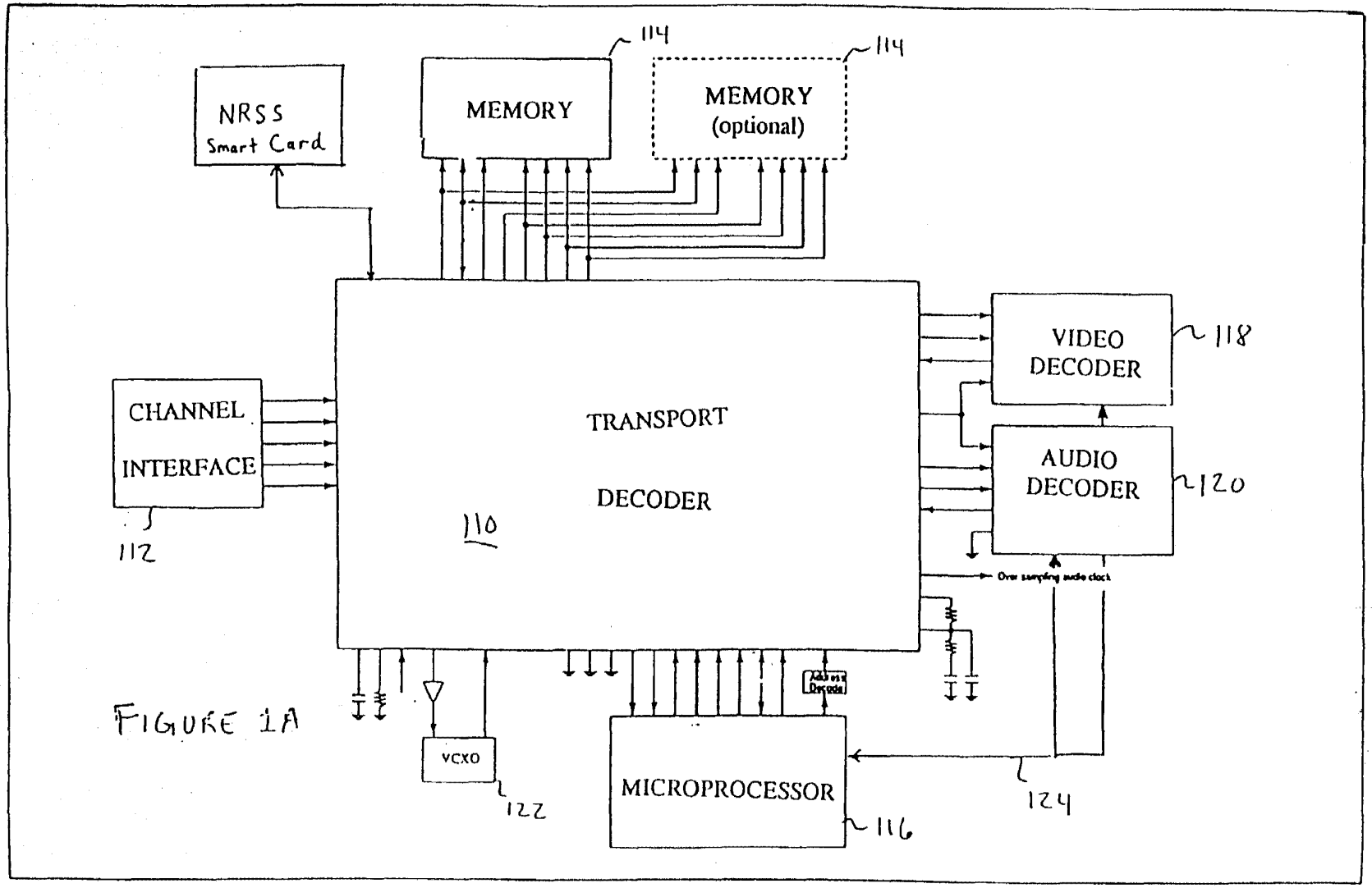

NRSS
Smart Card
MEMORY
114
MEMORY
(optional)
114
CHANNEL
INTERFACE
112
TRANSPORT
DECODER
110
VIDEO
DECODER
118
AUDIO
DECODER
120
Over sampling audio clock
VCXO
122
MICROPROCESSOR
116
124

FIGURE 1A

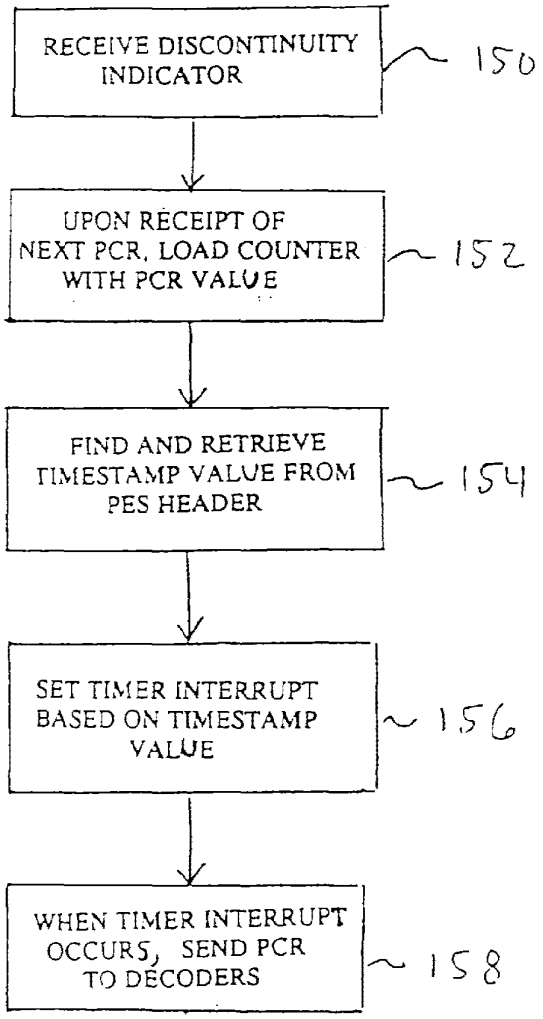

RECEIVE DISCONTINUITY INDICATOR
150
UPON RECEIPT OF NEXT PCR, LOAD COUNTER WITH PCR VALUE
152
FIND AND RETRIEVE TIMESTAMP VALUE FROM PES HEADER
154
SET TIMER INTERRUPT BASED ON TIMESTAMP VALUE
156
WHEN TIMER INTERRUPT OCCURS, SEND PCR TO DECODERS
158

FIG. 1B

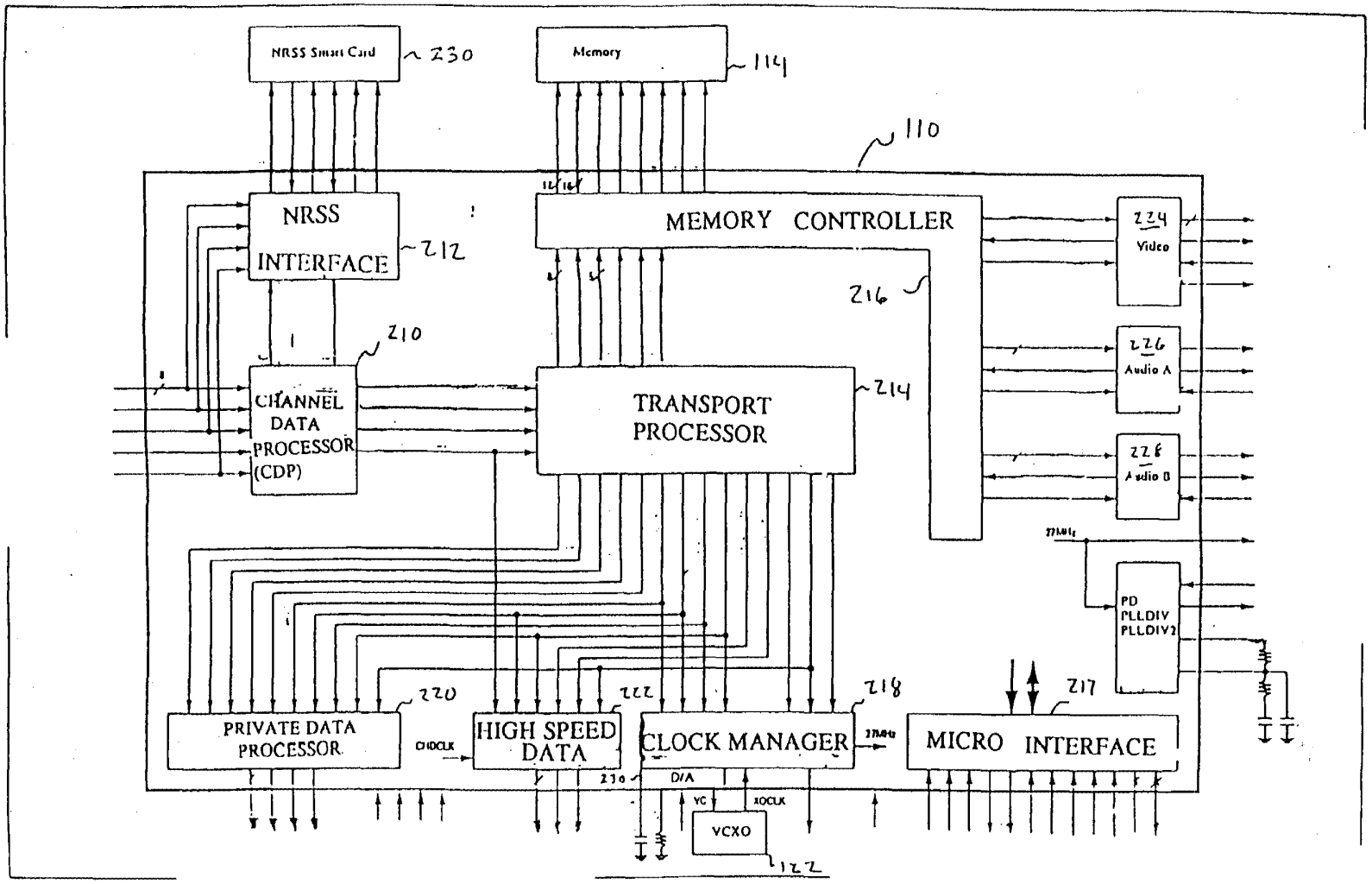

NRSS Smart Card
230
Memory
114
110
NRSS INTERFACE
212
MEMORY CONTROLLER
216
224
Video
210
226
Audio A
CHANNEL DATA PROCESSOR (CDP)
TRANSPORT PROCESSOR
214
228
Audio B
PD
PLLDIV
PLLDIV2
220
222
218
217
PRIVATE DATA PROCESSOR
CHDCLK
HIGH SPEED DATA
CLOCK MANAGER
MICRO INTERFACE
D/A
VC
XOCLK
VCXO
122

FIGURE 2A

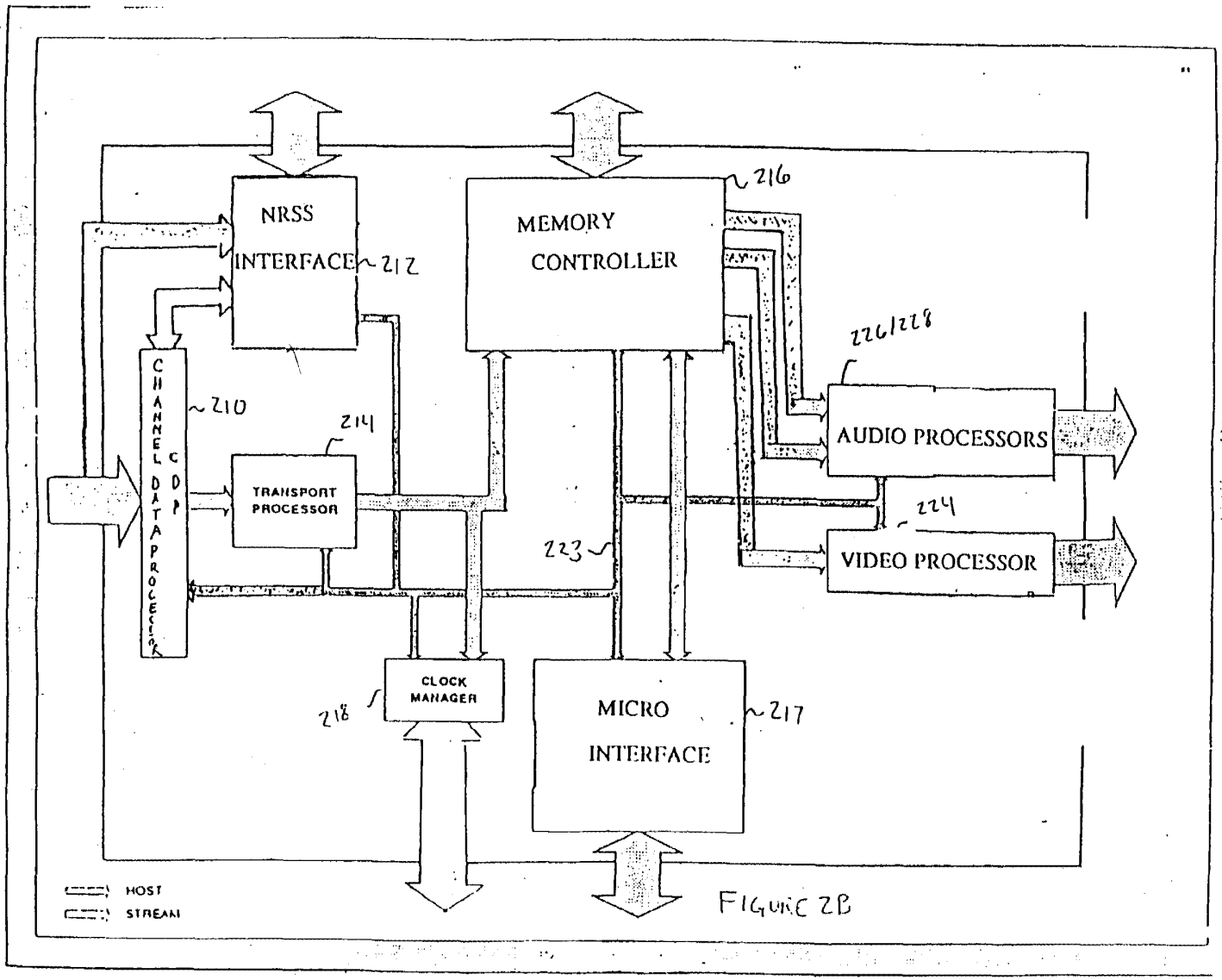

NRSS
INTERFACE
212
MEMORY
CONTROLLER
216
226/228
AUDIO PROCESSORS
224
VIDEO PROCESSOR
223
CHANNEL DATA PROCESSOR
210
CDP
214
TRANSPORT
PROCESSOR
CLOCK
MANAGER
218
MICRO
INTERFACE
217
HOST
STREAM

FIGURE 2B

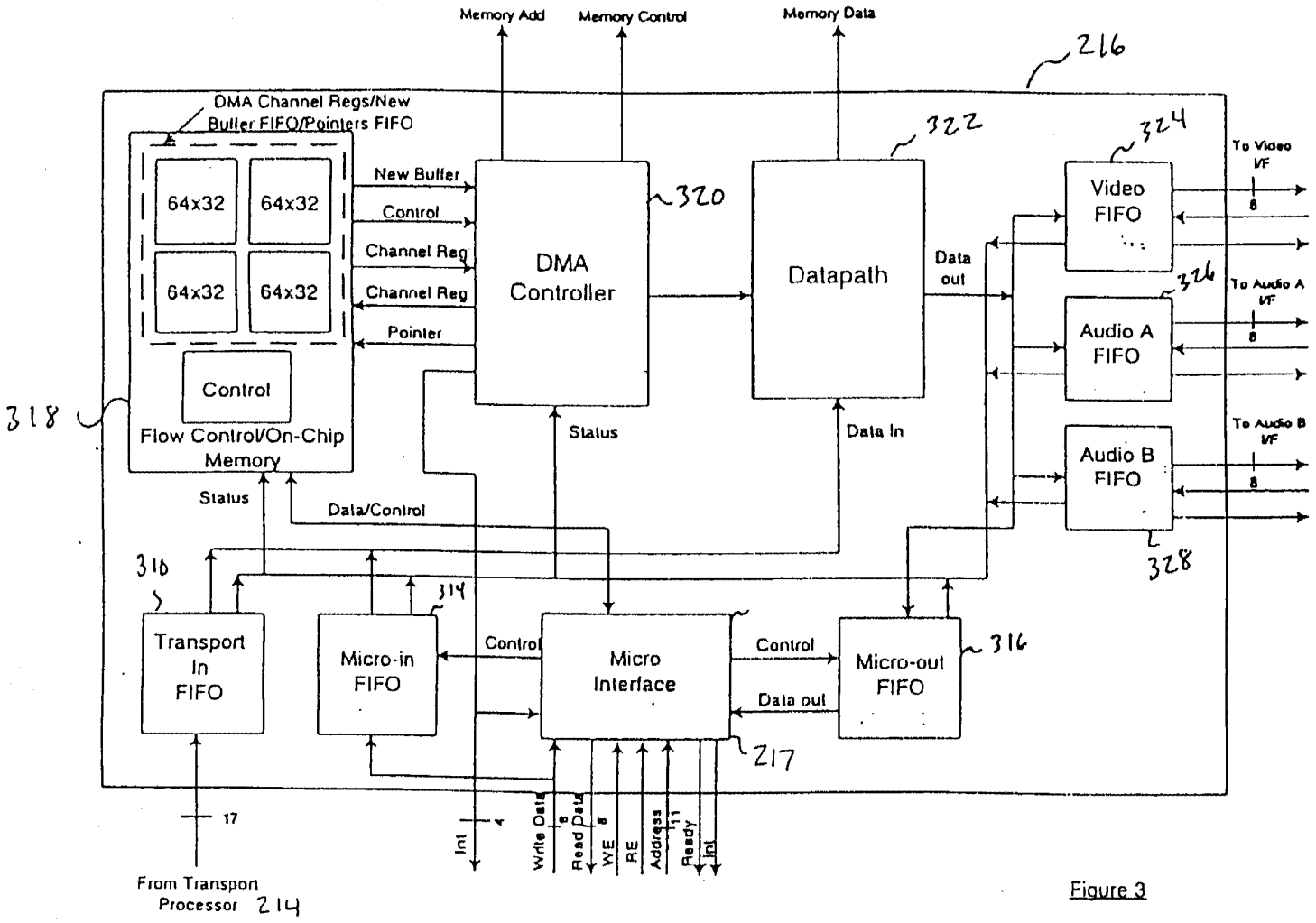
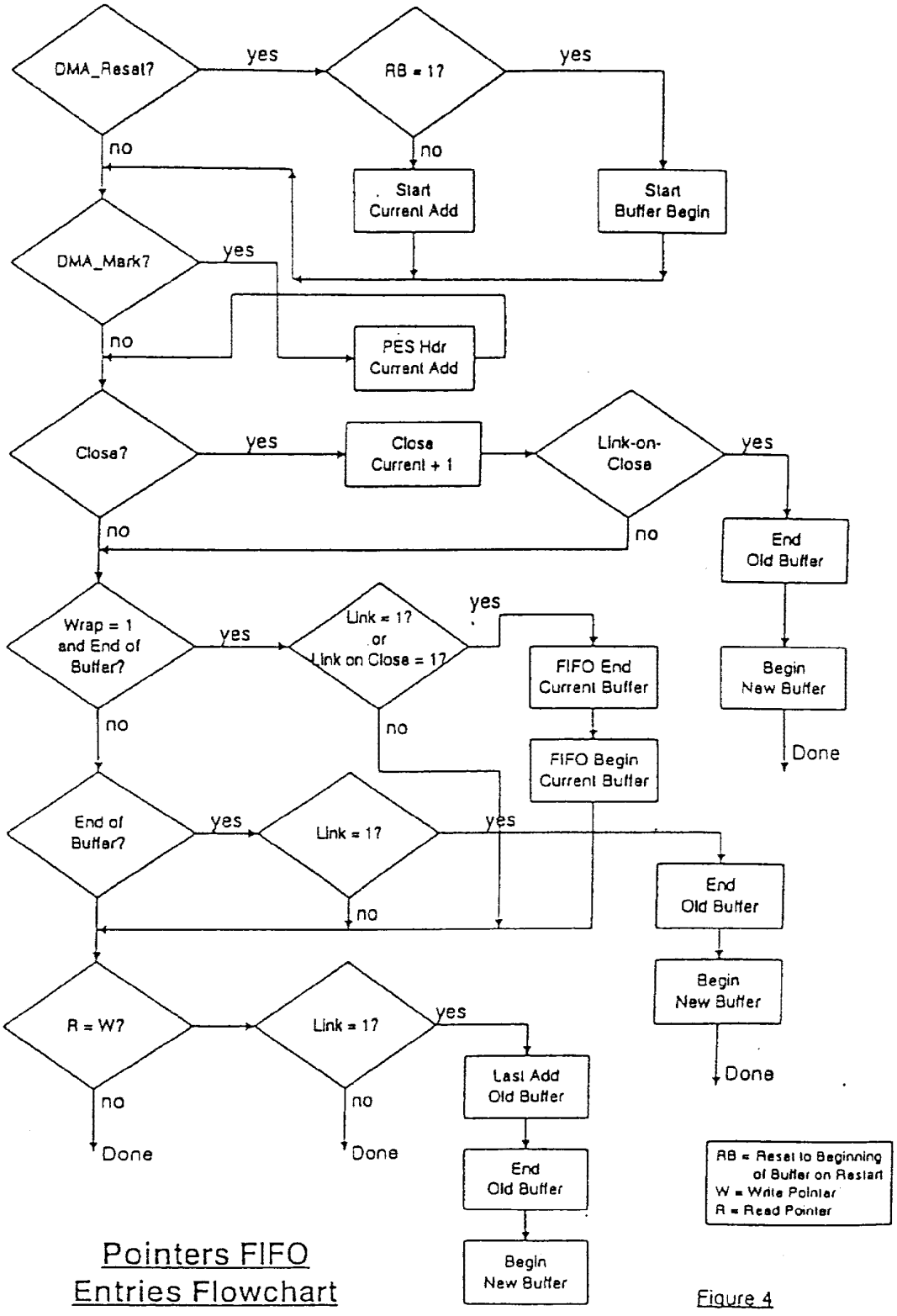


Figure 3

Memory Controller Block



Pointers FIFO Entries Flowchart

Figure 4

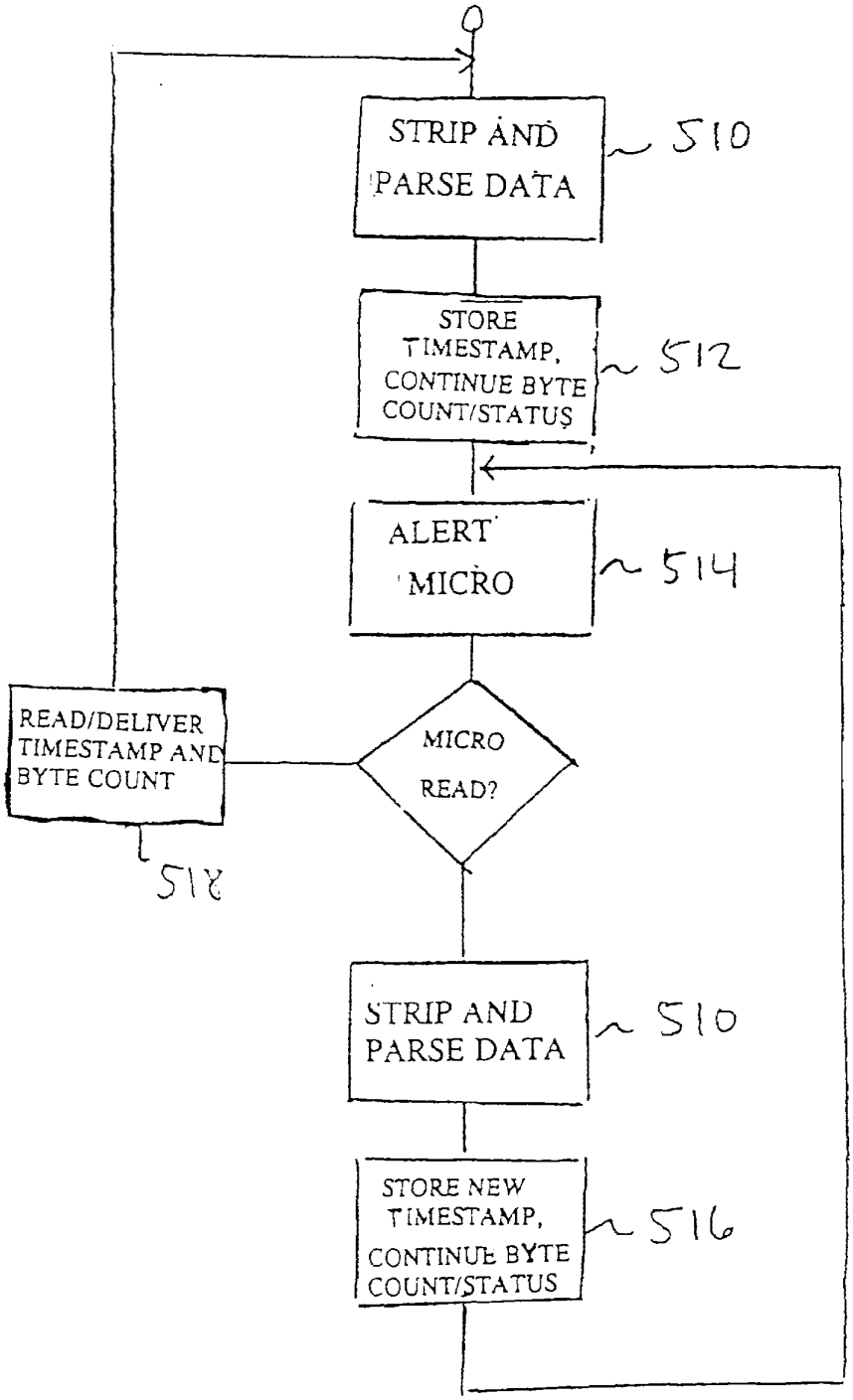

STRIP AND PARSE DATA
510
STORE TIMESTAMP, CONTINUE BYTE COUNT/STATUS
512
ALERT MICRO
514
MICRO READ?
READ/DELIVER TIMESTAMP AND BYTE COUNT
518
STRIP AND PARSE DATA
510
STORE NEW TIMESTAMP, CONTINUE BYTE COUNT/STATUS
516

FIGURE 5A

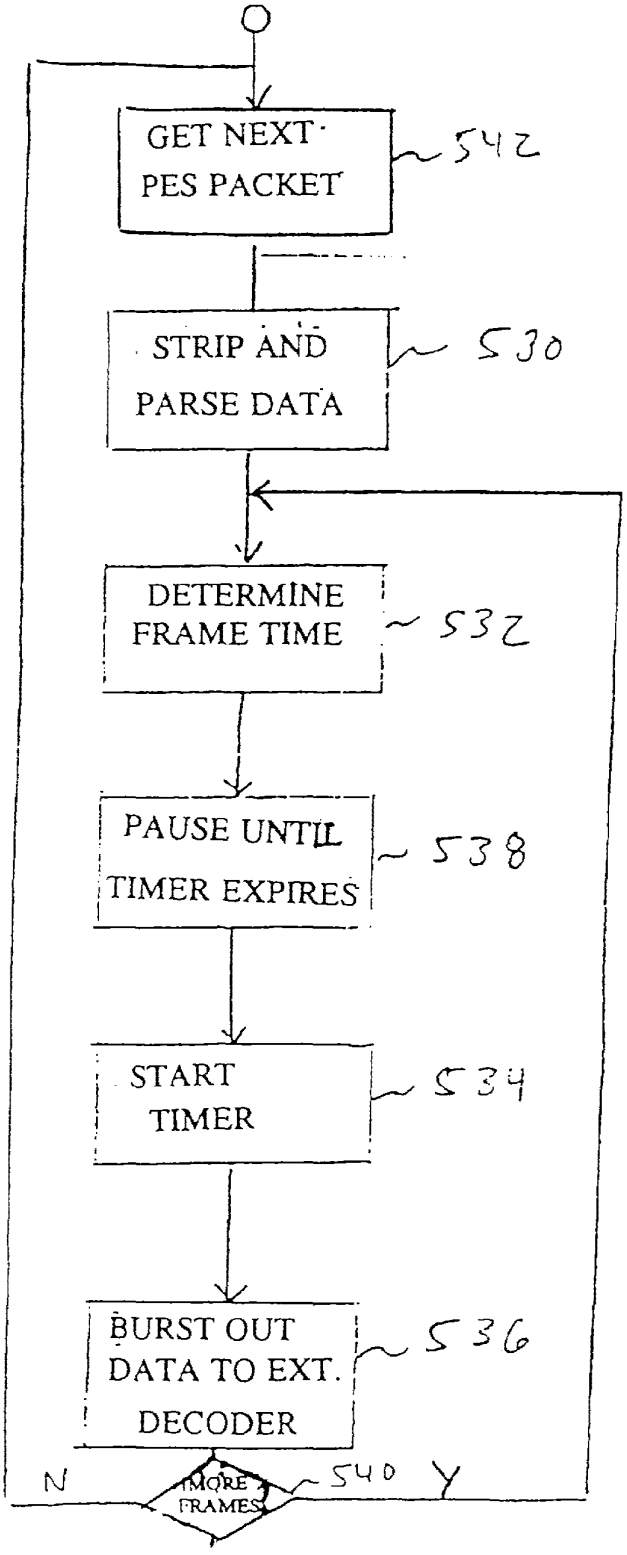

GET NEXT PES PACKET
542
STRIP AND PARSE DATA
530
DETERMINE FRAME TIME
532
PAUSE UNTIL TIMER EXPIRES
538
START TIMER
534
BURST OUT DATA TO EXT. DECODER
536
MORE FRAMES
540
N
Y

FIGURE 5B

Electronic Acknowledgement Receipt

EFS ID:	4945520
Application Number:	11338934
International Application Number:	
Confirmation Number:	6610
Title of Invention:	Grouping of image frames in video coding
First Named Inventor/Applicant Name:	Miska Hannuksela
Customer Number:	76385
Filer:	Erin M. Nichols/Tracey Dotter
Filer Authorized By:	Erin M. Nichols
Attorney Docket Number:	TPO.004.C2
Receipt Date:	12-MAR-2009
Filing Date:	25-JAN-2006
Time Stamp:	12:21:57
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Information Disclosure Statement Letter	IDS-5-TPO004C2.pdf	56138 998785e1d84e98bdca6f61f97d7a14ae6f0c0a16	no	2

Warnings:	
Information:	

2	Information Disclosure Statement (IDS) Filed (SB/08)	1449-5-TPO004C2.pdf	55772 17cd1e2281a4e7c7994d6617c0338ffa16f60597	no	1
Warnings:					
Information:					
This is not an USPTO supplied IDS fillable form					
3	NPL Documents	ChineseOfficeAction-TPO004C1.pdf	577884 a243ee21afd09a7477edd47fbf63fa7d7e8dafe3	no	9
Warnings:					
Information:					
4	Foreign Reference	CN1173783A.pdf	1780503 2bbcb9681463a00d81a20d40bac1322679baf26c	no	36
Warnings:					
Information:					
5	Foreign Reference	EP0798932.pdf	1101670 d7b7193a23b7a6f4e874c79af8c28b44b24b306d	no	22
Warnings:					
Information:					
		Total Files Size (in bytes):	3571967		

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

<u>New Applications Under 35 U.S.C. 111</u>
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

<u>National Stage of an International Application under 35 U.S.C. 371</u>
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

<u>New International Application Filed with the USPTO as a Receiving Office</u>
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

FORM 1449* INFORMATION DISCLOSURE STATEMENT IN AN APPLICATION (Use several sheets if necessary)	Docket Number: TPO.004.C2	Application Number: 11/338,934
	Applicant: HANNUKSELA	
	Filing Date: Jan. 25, 2006	Group Art Unit: 2613

U.S. PATENT DOCUMENTS

EXAMINER INITIAL	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	6,26,158	7/24/2001	Hata et al.			

FOREIGN PATENT DOCUMENTS

	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	TRANSLATION	
						YES	NO
	10-271511	10/09/1998	JP				X

OTHER DOCUMENTS (Including Author, Title, Date, Pertinent Pages, Etc.)

EXAMINER	DATE CONSIDERED
EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609; draw line through citation if not in conformance and not considered. Include copy of this form for next communication to the Applicant.	

*Substitute Disclosure Statement Form (PTO-1449) Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Electronic Acknowledgement Receipt

EFS ID:	5179746
Application Number:	11338934
International Application Number:	
Confirmation Number:	6610
Title of Invention:	Grouping of image frames in video coding
First Named Inventor/Applicant Name:	Miska Hannuksela
Customer Number:	76385
Filer:	Erin M. Nichols
Filer Authorized By:	
Attorney Docket Number:	TPO.004.C2
Receipt Date:	20-APR-2009
Filing Date:	25-JAN-2006
Time Stamp:	16:27:02
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Foreign Reference	JP10271511.pdf	1504119 30117798bbfb84b3a6032e8f31dcb02fa3fee47f	no	36

Warnings:	
Information:	

2	Transmittal Letter	IDS-6-TPO004C2.pdf	55984 8475d972cde663804faa27ebac7e402f8151d1ce	no	2
Warnings:					
Information:					
3	Information Disclosure Statement (IDS) Filed (SB/08)	Form1449supp6-TPO004C2.pdf	55814 d995fc2cc09eb5bb5a27f768a3556f275310f1a2	no	1
Warnings:					
Information:					
This is not an USPTO supplied IDS fillable form					
4	NPL Documents	JP-OfficeAction-26032009.pdf	185000 3510bcc0224e2113b3825c248487950a40000166	no	6
Warnings:					
Information:					
5	NPL Documents	JP-OfficeActionEngligh-26032009pdf.pdf	189167 c8af3ce98c9773d06b6f6a1077e25746014439df	no	9
Warnings:					
Information:					
		Total Files Size (in bytes):	1990084		

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

<u>New Applications Under 35 U.S.C. 111</u>
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

<u>National Stage of an International Application under 35 U.S.C. 371</u>
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

<u>New International Application Filed with the USPTO as a Receiving Office</u>
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

(19)日本国特許庁（ＪＰ）　(12) 公開特許公報（Ａ）　(11)特許出願公開番号

特開平10－271511

(43)公開日　平成10年(1998)10月９日

(51)Int.Cl.[6]	識別記号	ＦＩ
Ｈ０４Ｎ　7/32		Ｈ０４Ｎ　7/137

審査請求　未請求　請求項の数９　ＯＬ　（全 36 頁）

(21)出願番号　特願平9－130614

(22)出願日　平成９年(1997)５月21日

(31)優先権主張番号　特願平9－9369

(32)優先日　平９(1997)１月22日

(33)優先権主張国　日本（ＪＰ）

(71)出願人　000005821
松下電器産業株式会社
大阪府門真市大字門真1006番地

(72)発明者　畑　幸一
大阪府門真市大字門真1006番地　松下電器産業株式会社内

(72)発明者　栄藤　稔
大阪府門真市大字門真1006番地　松下電器産業株式会社内

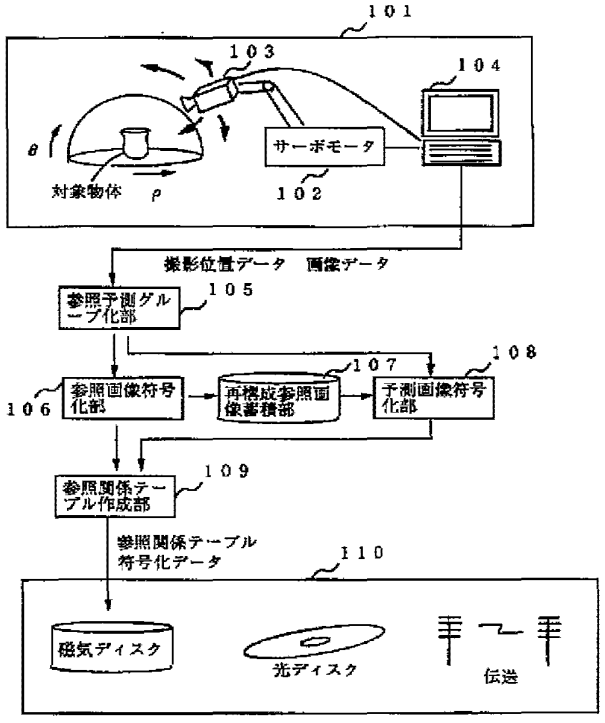
(74)代理人　弁理士　滝本　智之　（外１名）

(54)【発明の名称】　画像符号化装置と画像復号化装置

(57)【要約】

【課題】　被符号化画像群の符号化データに予測符号化の際の参照関係と撮影位置と符号化データ位置のテーブルを付加することで、効率のよく予測符号化でき、要求に効率よく対応できる復号化を実現する。

【解決手段】　複数の被符号化画像とその撮影位置データを入力とし、前記複数の被符号化画像を参照画像と参照画像から予測符号化される予測画像にグループ化する参照予測グループ化部１０５と、参照画像を符号化する参照画像符号化部１０６と、再構成参照画像を用いて予測画像を符号化する予測画像符号化部１０８と、参照関係と符号化データ位置と撮影位置を記述した参照関係テーブルを作成し、参照関係テーブルと符号化データを出力する参照関係テーブル作成部１０９手段を備えたことを特徴とする。



1

【特許請求の範囲】

【請求項１】複数の被符号化画像とその撮影位置データを入力とし、前記複数の被符号化画像を参照画像と参照画像から予測符号化される予測画像にグループ化する参照予測グループ化手段と、参照画像を符号化する参照画像符号化手段と、再構成参照画像を用いて予測画像を符号化する予測画像符号化手段と、参照関係と符号化データ位置と撮影位置を記述した参照関係テーブルを作成し、参照関係テーブルと符号化データを出力する参照関係テーブル作成手段を備えたことを特徴とする画像符号化装置。

【請求項２】再生すべき画像の要求と、画像が符号化された符号化データと、参照予測関係と撮影位置と符号化データ位置を記述した参照関係テーブルを入力とし、参照関係テーブルから要求される画像の復号化に必要な符号化データを選択する符号化データ選択手段と、符号化データを復号化し、再構成参照画像を出力する参照画像復号化手段と、符号化データを復号化し再構成予測画像を出力する予測画像復号化手段を備えたことを特徴とする画像復号化装置。

【請求項３】撮影位置と、参照画像符号化データ位置と予測画像符号化データ位置を記述したデータ位置テーブルと、予測画像符号化データについては、参照画像に対する変形情報とこの変形情報を用いて変形された画像に対する残差情報で構成されることを特徴とする符号化データを入力とし、再生すべき画像要求を逐次監視する画像要求監視手段と、画像要求と画像復号化状況とデータ位置テーブルと符号化データから、復号化する符号化データを選択する符号化データ選択手段と、符号化データから再構成参照画像を復号化する再構成参照画像復号化手段と、再構成参照画像と符号化データの変形情報から変形参照画像を復号化する変形参照画像復号化手段と、変形画像と符号化データの残差情報から再構成予測画像を復号化する再構成予測画像復号化手段を備え、再構成参照画像復号化手段と変形参照画像復号化手段および再構成予測画像復号化手段の出力を選択的に出力することを特徴とする画像復号化装置。

【請求項４】再生すべき画像の要求と、撮影位置と参照画像符号化データ位置と予測画像符号化データ位置を記述したデータ位置テーブルと、画像が符号化された符号化データを入力とし、復号化する符号化データを選択する符号化データ選択手段と、符号化データから参照画像を復号化する参照画像復号化手段と、再構成された参照画像を複数枚蓄積する再構成参照画像群蓄積手段と、符号化データから予測画像を復号化する予測画像復号化手段と、再構成された予測画像を複数枚蓄積する再構成予測画像群蓄積手段と、再生すべき画像の要求に応じて蓄積すべき画像の撮影位置を計算し、前記再構成参照画像群蓄積手段と再構成予測画像群蓄積手段に蓄積されている画像量を一定量以下に保つ画像蓄積検査手段を備え、

2

前記復号処理済みの蓄積されている画像から再生画像を出力することを特徴とする画像復号化装置。

【請求項５】撮影位置と参照画像符号化データ位置と予測画像符号化データ位置を記述したデータ位置テーブルと、符号化データを入力とし、再生すべき画像の要求を逐次監視し、画像要求を出力する画像要求監視手段と、現在の画像要求と前回の画像要求と既に復号化され蓄積された画像の情報から、新たに復号化する符号化データを選択する符号化データ選択手段と、撮影位置に関して現在の画像要求近傍の画像の符号化データを選択する予測符号化データ選択手段と、符号化データを参照画像に復号化する参照画像復号化手段と、符号化データを予測画像に復号化する予測画像復号化手段と、復号化された参照画像を蓄積する再構成参照画像群蓄積手段と、復号化された予測画像を蓄積する再構成予測画像群蓄積手段とを有し、画像要求を監視し再生する一方で、現在の画像要求近傍の画像の符号化データを選択し復号を行なうことを特徴とする画像復号化装置。

【請求項６】複数の被符号化画像と画像撮影位置データを入力とし、前記複数の被符号化画像を参照画像と、参照画像から予測される予測画像にグループ化する参照予測グループ化手段と、参照画像と予測画像の間の変形拘束式を推定する変形拘束式推定手段と、参照画像を符号化する参照画像符号化手段と、変形拘束式に基づいて変形を拘束して予測画像を符号化する拘束予測画像符号化手段、参照予測グループの参照関係と符号化データ位置を記述した参照関係テーブルを作成し、参照関係テーブルと符号化データを出力する参照関係テーブル作成手段を備えたことを特徴とする画像符号化装置。

【請求項７】再生すべき画像の要求と、画像が符号化された符号化データと、参照予測関係と撮影位置と符号化データ位置を記述した参照関係テーブルを入力とし、参照関係テーブルから要求される画像の復号化に必要な符号化データを選択する符号化データ選択手段と、参照画像符号化データを復号化し、再構成参照画像を出力する参照画像復号化手段と、変形拘束式に基づいて変形を拘束して予測画像符号化データを復号化し再構成予測画像を出力する拘束予測画像復号化手段を備えたことを特徴とする画像復号化装置。

【請求項８】複数の被符号化画像とその撮影位置データと光源位置データを入力とし、前記複数の被符号化画像を参照画像と参照画像から予測符号化される予測画像にグループ化する参照予測グループ化手段と、参照画像を符号化する参照画像符号化手段と、再構成参照画像を用いて予測画像を符号化する予測画像符号化手段と、参照関係と符号化データ位置と撮影位置を記述した参照関係テーブルを作成し、参照関係テーブルと符号化データを出力する参照関係テーブル作成手段を備えたことを特徴とする画像符号化装置。

【請求項９】再生すべき画像の要求と、画像が符号化さ

れた符号化データと、参照予測関係と撮影位置と光源位置と符号化データ位置を記述した参照関係テーブルを入力とし、参照関係テーブルから要求される画像の復号化に必要な符号化データを選択する符号化データ選択手段と、符号化データを復号化し、再構成参照画像を出力する参照画像復号化手段と、符号化データを復号化し再構成予測画像を出力する予測画像復号化手段を備えたことを特徴とする画像復号化装置。

【発明の詳細な説明】

【０００１】

【発明の属する技術分野】本発明は画像の伝送・蓄積のための画像符号化装置と画像復号化装置に関するものである。

【０００２】

【従来の技術】従来、３次元物体を計算機上で表現する際、物体の形状情報や表面の材質情報を計算機に入力し、見え方を計算し、シュミレーションする手法が取られているが、複雑な形状や細かな質感を表現するには、多くのデータを入力する必要があり、シュミレーションの計算量も膨大なものとなる。

【０００３】一方、より小さい計算量で３次元物体を表現するために、イメージベースドレンダリング（Image-BasedRendering）と呼ばれる、多くの視点からの画像をあらかじめ撮影・蓄積・伝送しておき、それらを要求に応じて表示する手法も取られている（「クイックタイムブイアール」（S. E. Chen、 "QuickTime VR — Animage-Based Approach to Virtual Environment Navigation"、 SIGGRAPH'90、 pp. 29—38、 1990））。

【０００４】

【発明が解決しようとする課題】Image-Based Rendering の手法では、多くの要求に応えるために、多くの視点からの画像をあらかじめ撮影しておくため、膨大なデータ量を蓄積・伝送しなければならない、これらのデータは画像として約１万枚、約２．４ギガバイトを越えることもある。したがって、これらの膨大なデータ量を、効率的に蓄積・伝送することが課題となっている。

【０００５】本発明は、従来のこのような課題を考慮し、蓄積・伝送すべき多数の画像の中から少数の代表画像を選び、各代表画像で複数枚の画像を表現し、その依存関係を付加することによって、効率的に複数の画像を符号化する画像符号化装置や、段階的な復号化や、次の要求の予測を行うことによって効率的に画像を符号化または復号化する方式とその装置を提供することを目的とする。

【０００６】

【課題を解決するための手段】第１の発明は、複数の被符号化画像とその撮影位置データを入力とし、前記複数の被符号化画像を参照画像と参照画像から予測符号化される予測画像にグループ化する参照予測グループ化手段と、参照画像を符号化する参照画像符号化手段と、再構

成参照画像を用いて予測画像を符号化する予測画像符号化手段と、参照関係と符号化データ位置と撮影位置を記述した参照関係テーブルを作成し、参照関係テーブルと符号化データを出力する参照関係テーブル作成手段を備えたことを特徴とする画像符号化装置である。

【０００７】第２の発明は、再生すべき画像の要求と、画像が符号化された符号化データと、参照予測関係と撮影位置と符号化データ位置を記述した参照関係テーブルを入力とし、参照関係テーブルから要求される画像の復号化に必要な符号化データを選択する符号化データ選択手段と、符号化データを復号化し、再構成参照画像を出力する参照画像復号化手段と、符号化データを復号化し再構成予測画像を出力する予測画像復号化手段を備えたことを特徴とする画像復号化装置である。

【０００８】第３の発明は撮影位置と、参照画像符号化データ位置と予測画像符号化データ位置を記述したデータ位置テーブルと、予測画像符号化データについては、参照画像に対する変形情報とこの変形情報を用いて変形された画像に対する残差情報で構成されることを特徴とする符号化データを入力とし、再生すべき画像要求を逐次監視する画像要求監視手段と、画像要求と画像復号化状況とデータ位置テーブルと符号化データから、復号化する符号化データを選択する符号化データ選択手段と、符号化データから再構成参照画像を復号化する再構成参照画像復号化手段と、再構成参照画像と符号化データの変形情報から変形参照画像を復号化する変形参照画像復号化手段と、変形画像と符号化データの残差情報から再構成予測画像を復号化する再構成予測画像復号化手段を備え、再構成参照画像復号化手段と変形参照画像復号化手段および再構成予測画像復号化手段の出力を選択的に出力することを特徴とする画像復号化装置である。

【０００９】第４の発明は、再生すべき画像の要求と、撮影位置と参照画像符号化データ位置と予測画像符号化データ位置を記述したデータ位置テーブルと、画像が符号化された符号化データを入力とし、復号化する符号化データを選択する符号化データ選択手段と、符号化データから参照画像を復号化する参照画像復号化手段と、再構成された参照画像を複数枚蓄積する再構成参照画像群蓄積手段と、符号化データから予測画像を復号化する予測画像復号化手段と、再構成された予測画像を複数枚蓄積する再構成予測画像群蓄積手段と、再生すべき画像の要求に応じて蓄積すべき画像の撮影位置を計算し、前記再構成参照画像群蓄積手段と再構成予測画像群蓄積手段に蓄積されている画像量を一定量以下に保つ画像蓄積検査手段を備え、前記復号処理済みの蓄積されている画像から再生画像を出力することを特徴とする画像復号化装置である。

【００１０】第５の発明は、撮影位置と参照画像符号化データ位置と予測画像符号化データ位置を記述したデータ位置テーブルと、符号化データを入力とし、再生すべ

き画像の要求を逐次監視し、画像要求を出力する画像要求監視手段と、現在の画像要求と前回の画像要求と既に復号化され蓄積された画像の情報から、新たに復号化する符号化データを選択する符号化データ選択手段と、撮影位置に関して現在の画像要求近傍の画像の符号化データを選択する予測符号化データ選択手段と、符号化データを参照画像に復号化する参照画像復号化手段と、符号化データを予測画像に復号化する予測画像復号化手段と、復号化された参照画像を蓄積する再構成参照画像群蓄積手段と、復号化された予測画像を蓄積する再構成予測画像群蓄積手段とを有し、画像要求を監視し再生する一方で、現在の画像要求近傍の画像の符号化データを選択し復号を行なうことを特徴とする画像復号化装置である。

【００１１】第６の発明は、複数の被符号化画像と画像撮影位置データを入力とし、前記複数の被符号化画像を参照画像と、参照画像から予測される予測画像にグループ化し、参照画像と予測画像の間の変形拘束式を推定し、参照画像を符号化し、変形拘束式に基づいて変形を拘束して予測画像を符号化し、参照予測グループの参照関係と符号化データ位置を記述した参照関係テーブルを作成し、参照関係テーブルと符号化データを出力することを特徴とする画像符号化装置である。

【００１２】第７の発明は、再生すべき画像の要求と、画像が符号化された符号化データと、参照予測関係と撮影位置と符号化データ位置を記述した参照関係テーブルを入力とし、参照関係テーブルから要求される画像の復号化に必要な符号化データを選択し、参照画像符号化データを復号化し、再構成参照画像を出力し、変形拘束式に基づいて変形を拘束して予測画像符号化データを復号化し再構成予測画像を出力することを特徴とする画像復号化装置である。

【００１３】第８の発明は、複数の被符号化画像とその撮影位置データと光源位置データを入力とし、前記複数の被符号化画像を参照画像と参照画像から予測符号化される予測画像にグループ化し、参照画像を符号化し、再構成参照画像を用いて予測画像を符号化し、参照関係と符号化データ位置と撮影位置を記述した参照関係テーブルを作成し、参照関係テーブルと符号化データを出力することを特徴とする画像符号化装置である。

【００１４】第９の発明は、再生すべき画像の要求と、画像が符号化された符号化データと、参照予測関係と撮影位置と光源位置と符号化データ位置を記述した参照関係テーブルを入力とし、参照関係テーブルから要求される画像の復号化に必要な符号化データを選択し、符号化データを復号化し再構成参照画像を出力し、符号化データを復号化し再構成予測画像を出力することを特徴とする画像復号化装置である。

【００１５】第１の発明に基づく画像符号化装置では、参照予測グループ化手段が、入力である複数の被符号化

画像とその撮影位置データから被符号化画像を参照画像と参照画像から予測符号化される予測画像からなる複数のグループに分ける。参照画像は参照画像符号化手段により符号化され、その再構成画像を用いて予測画像符号化手段が予測画像を符号化する。そして、参照関係テーブル作成手段が、参照画像に対する予測画像の参照関係、符号化データ位置、撮影位置を記述した参照関係テーブルを作成し、参照関係テーブルと符号化データを出力する。

【００１６】第２の発明に基づく画像復号化装置では、再生すべき画像の要求、画像符号化データおよび参照予測関係、撮影位置、符号化データ位置を記述した参照関係テーブルを入力とする。符号化データ選択手段は、要求される画像の復号化に必要な符号化データを参照関係テーブルに基づいて選択する。

【００１７】選択された符号化データが参照画像であれば、参照画像復号化手段により復号化され、再構成参照画像として出力される。また選択された符号化データが予測画像データであれば、予測画像復号化手段により復号化され再構成予測画像として出力される。

【００１８】第３の発明に基づく画像復号化装置では、撮影位置と参照画像符号化データ位置と予測画像符号化データ位置を記述したデータ位置テーブルと、予測画像符号化データについては、参照画像に対する変形情報とこの変形情報を用いて変形された画像に対する残差情報で構成されることを特徴とする符号化データを入力とする。画像要求監視手段は再生すべき画像要求を逐次監視する。符号化データ選択手段は、画像要求と画像復号化状況に応じてデータ位置テーブルと符号化データから、復号化すべき符号化データを選択する。

【００１９】選択された符号化データが参照画像データであれば、再構成参照画像復号化手段がその符号化データから再構成参照画像を復号化する。また選択された符号化データが予測画像データであれば、変形参照画像復号化手段または再構成予測画像復号化手段が予測画像データを復号する。変形参照画像復号化手段によれば、再構成参照画像と符号化データの変形情報から変形参照画像が復元され、再構成予測画像復号化手段によれば、変形画像と符号化データの残差情報から再構成予測画像が復元される。再構成参照画像復号化手段と変形参照画像復号化手段および再構成予測画像復号化手段の出力は、画像要求と画像復号化状況に応じてデータ位置テーブルにより選択的に出力されることになる。

【００２０】第４の発明に基づく画像復号化装置では、再生すべき画像の要求と、撮影位置と参照画像符号化データ位置と予測画像符号化データ位置を記述したデータ位置テーブルと、画像が符号化された符号化データを入力とする。

【００２１】符号化データ選択手段は再生すべき画像の要求を入力として復号化すべき符号化データを選択する。

選択された符号化データが参照画像データであれば、参照画像復号化手段が、選択された符号化データが予測画像データであれば、予測画像復号化手段が画像の再構成を行ない、その結果はそれぞれ再構成参照画像群蓄積手段と再構成予測画像群蓄積手段に蓄積される。この時、画像蓄積検査手段が、再生すべき画像の要求に応じて蓄積すべき画像の撮影位置を計算し、前記再構成参照画像群蓄積手段と再構成予測画像群蓄積手段に蓄積されている画像量を一定量以下に保つよう廃棄する蓄積画像を決定する。出力画像は復号処理済みの蓄積されている画像から選択されて出力される。

【００２２】第５の発明に基づく画像復号化装置では、撮影位置と参照画像符号化データ位置と予測画像符号化データ位置を記述したデータ位置テーブルと、符号化データを入力とする。画像要求監視手段は、再生すべき画像の要求を逐次監視し、画像要求を出力する。符号化データ選択手段は、現在の画像要求と前回の画像要求と既に復号化され蓄積された画像の情報から、新たに復号化する符号化データを選択する。その一方で、予測符号化データ選択手段は、撮影位置に関して現在の画像要求近傍の画像の符号化データを選択する。以上の二つの選択手段で選ばれた符号化データは選択された符号化データが参照画像データであれば、参照画像復号化手段が、選択された符号化データが予測画像データであれば、予測画像復号化手段が画像の再構成を行ない、その結果はそれぞれ再構成参照画像群蓄積手段と再構成予測画像群蓄積手段に蓄積される。

【００２３】このように、画像要求を監視し再生する一方で、現在の画像要求近傍の画像の符号化データを選択し復号を行ない、画像を出力する。

【００２４】第６の発明に基づく画像符号化装置は、複数の被符号化画像と画像撮影位置データを入力とする。参照予測グループ化手段では、前記複数の被符号化画像を、参照画像と、参照画像から予測される予測画像にグループ化する。変形拘束式推定手段では、参照画像と予測画像の間の変形拘束式を推定する。

【００２５】参照画像符号化手段では、参照画像を符号化する。拘束予測画像符号化手段では、変形拘束式に基づいて変形を拘束して予測画像を符号化する。参照関係テーブル作成手段では、参照予測グループの参照関係と符号化データ位置を記述した参照関係テーブルを作成し、参照関係テーブルと符号化データを出力する。

【００２６】第７の発明に基づく画像復号化装置では、再生すべき画像の要求と、画像が符号化された符号化データと、参照予測関係と撮影位置と符号化データ位置を記述した参照関係テーブルを入力とする。符号化データ選択手段では、参照関係テーブルから要求される画像の復号化に必要な符号化データを選択する。参照画像復号化手段では、参照画像符号化データを復号化し、再構成参照画像を出力する。拘束予測画像復号化手段では、変

形拘束式に基づいて変形を拘束して予測画像符号化データを復号化し再構成予測画像を出力する。

【００２７】第８の発明に基づく画像符号化装置では、複数の被符号化画像とその撮影位置データと光源位置データを入力とする。参照予測グループ化手段では、前記複数の被符号化画像を参照画像と参照画像から予測符号化される予測画像にグループ化する。参照画像符号化手段では、参照画像を符号化する。予測画像符号化手段では、再構成参照画像を用いて予測画像を符号化する。参照関係テーブル作成手段では、参照関係と符号化データ位置と撮影位置を記述した参照関係テーブルを作成し、参照関係テーブルと符号化データを出力する。

【００２８】第９の発明に基づく画像復号化装置では、再生すべき画像の要求と、画像が符号化された符号化データと、参照予測関係と撮影位置と光源位置と符号化データ位置を記述した参照関係テーブルを入力とする。

【００２９】符号化データ選択手段では、参照関係テーブルから要求される画像の復号化に必要な符号化データを選択する。参照画像復号化手段では、符号化データを復号化し、再構成参照画像を出力する。予測画像復号化手段では、符号化データを復号化し再構成予測画像を出力する。

【００３０】

【発明の実施の形態】以下、本発明にかかる実施の形態について図面を参照しながら説明する。

【００３１】図１は本発明の第１の実施の形態である画像符号化装置の構成を示すブロック図であり、同図を用いて本実施の形態の構成を説明する。

【００３２】同図において、撮影位置制御型画像撮像部（１０１）は、対象物体の画像を撮影し、その撮影位置データと画像データを出力する。参照予測グループ部（１０５）は、撮影位置データと画像データを参照画像と一枚以上の予測画像からなる参照予測グループにグループ化し、参照画像を参照画像符号化部（１０６）に出力し、予測画像を予測画像符号化部（１０８）に出力する。

【００３３】参照画像符号化部（１０６）は、参照画像を符号化し、再構成参照画像を作成し、符号化データを参照関係作成部（１０９）に出力し、再構成参照画像を再構成参照画像蓄積部（１０７）に出力する。再構成参照画像蓄積部（１０７）は、再構成参照画像を蓄積し、予測画像符号化部（１０８）に再構成参照画像を出力する。

【００３４】予測画像符号化部（１０８）は、予測画像を符号化し、符号化データを参照関係作成部（１０９）に出力する。参照関係テーブル作成部（１０９）は、符号化データから参照関係と符号化データ位置と撮影位置を記述した参照関係テーブルを作成し、参照関係テーブルと符号化データをデータ記録・伝送部（１１０）に出力する。データ記録・伝送部（１１０）は、用途に応じ

て記録・伝送する。

【００３５】以上のように構成された本実施の形態の画像符号化装置の動作を説明する。まず、撮影位置制御型画像撮像部（１０１）では、対象物体の周囲を網羅するように、異なる視点から画像を撮影する。計算機によってサーボモータ（１０２）を対象物体の正面から水平方向にρ度、水平面から垂直方向にθ度の位置にカメラ（１０３）を移動させ、撮影位置・画像データを計算機（１０４）によって取得する。画像データは、クロマキーによって対象物体を背景と分離し、透過度情報を示すアルファプレーンを付加し、輝度プレーン（２０１）、色差１プレーン（２０２）、色差２プレーン（２０３）、アルファプレーン（２０４）で構成する。

【００３６】ここで、ρ，θは、

【００３７】

【数１】

$$0 \leq \rho < 360,\ 0 \leq \theta \leq 90,\ (\rho,\theta\text{は整数})$$

【００３８】とする。このように取得された撮影位置・画像データは、図３に示すように、各画像データの先頭に水平方向撮影角度ρと、垂直方向撮影角度θを付加した書式で出力される。

【００３９】参照予測グループ化部（１０５）では、参照画像（４０１）によって類似した画像を予測画像（４０２）として予測符号化するために、$\rho-\theta$空間において＊て等間隔に参照画像を配置し、参照画像の近傍の画像をその参照画像より予測される予測画像にグループ化するため、撮影位置Ｐ＝（ρ，θ）を用いて、以下のように画像それぞれに参照画像か予測画像のフラグｆ（Ｐ）を割り当てる。

【００４０】

【数２】

$$f(\mathbf{P}) = \begin{cases} 0 & \mathbf{P}=(Sn,Sm)\ (n,m\text{は整数}) \\ 1 & \text{それ以外} \end{cases}$$

【００４１】さらに、以下のようにインデックスｎ（Ｐ）を割り当てる。

【００４２】

【数３】

$$n(\mathbf{P}) = \lceil \tfrac{359}{S} \rceil \times n + m,\ n,m = \min_{n,m} d(\mathbf{P},\mathbf{Q})$$

【００４３】以上において、Ｑ＝（Ｓｎ，Ｓｍ）である。Ｓは参照画像間の間隔でありＳ＝５である。「・」は、・を越えない整数を表わす。また、水平方向撮影位置で折り返すため、

【００４４】

【数４】

$$\mathbf{P}_1 = (\rho_1,\theta_1),\ \mathbf{P}_2 = (\rho_2,\theta_2)$$

【００４５】とすると、

【００４６】

【数５】

$$d(\mathbf{P}_1,\mathbf{P}_2) = \min((\rho_1-\rho_2)^2,(\rho_1-\rho_2-360)^2,(\rho_1-\rho_2+360)^2) + (\theta_1-\theta_2)^2$$

【００４７】である。これらを、インデックス順に図５の示す書式で、参照画像は参照画像符号化部（１０６）へ、予測画像は予測画像符号化部（１０７）へ出力する。

【００４８】参照画像符号化部（１０６）では、参照画像をその画像データのみで符号化もしくは、その符号化データのみで復号化出来るよう復号化しておく、アルファプレーンを含んだ、そのような符号化する方法には、ＭＰＥＧ４（MPEG Video Group、"MPEG-4 Video Verification Model Version 5.0"、ISO/IEC JTC1/SC29/WG11 Document N1469、 1996）で標準化が進められている動画像符号化方法のＩ－ＰＩＣＴＵＲＥ符号化方式がある。

【００４９】この方法を用いて、参照画像符号化部（１０６）で参照画像を符号化し、符号化データを参照関係テーブル作成部（１０９）に出力し、参照画像の再構成画像を作成し、再構成参照画像蓄積部（１０７）に出力する。図６に参照画像符号化部（１０６）のフローチャートを示す。

【００５０】再構成参照画像蓄積部（１０７）では、再構成参照画像を蓄積しておき、予測画像符号化部（１０８）に出力する。参照予測グループ化部（１０５）において、図５のように参照画像と予測画像が整列されて出力されるので、再構成参照画像蓄積部（１０８）は、画像１枚分の容量を持っていれば十分である。

【００５１】予測画像符号化部（１０８）では、予測画像を再構成参照画像を用いて符号化する。参照画像と同様に、アルファプレーンを含んだ予測画像符号化方法には、ＭＰＥＧ４による動画像符号化のＰ－ＰＩＣＴＵＲＥ符号化方式がある。この方法を用いて予測画像を符号化し、符号化データを参照関係テーブル作成部（１０９）に出力する。図７に参照画像符号化部（１０６）のフローチャートを示す。

【００５２】参照関係テーブル作成部（１０９）では、各符号化部より出力された図８のようなデータを順に走査し、図９のような参照関係テーブルと図１０のような符号化データを作成し、データ記録・伝送部（１１０）に出力する。

【００５３】データ記録・伝送部（１１０）では、参照関係テーブルと符号化データを磁気ディスクや光ディスクに蓄積しておいたり、ネットワークなどで伝送する。

【００５４】以上のように、本実施の形態では、異なる多くの視点から撮影された多視点画像を、撮影位置データが近傍の類似した画像である参照画像と１枚以上の予測画像にグループ化し、効率よく符号化することが可能である。さらに、画像を各参照予測グループで参照画像が先頭になるように整列して、符号化手段に出力することで、計算量と計算に必要な蓄積容量を少なくしている。また、その整列順に符号化データ位置を格納したテーブルを付加することで、ランダムアクセスに対応した

効率的な復号化が可能である。

【００５５】図１１は本発明の第２の実施の形態である画像復号化装置の構成を示すブロック図であり、同図を用いて本実施の形態の構成を説明する。

【００５６】同図において、データ記憶・伝送部（１１０１）では、参照関係テーブルと符号化データを蓄積・伝送し、参照テーブルと符号化データを符号化データ選択部（１１０２）に出力する。符号化データ選択部（１１０２）では、参照関係テーブルからデータの位置を検索し、符号化データから必要なデータを参照画像復号化部（１１０３）や予測画像復号化部（１１０５）に出力する。参照画像復号化部（１１０３）では、参照画像符号化データを復号化し、再構成参照画像を再構成参照画像蓄積部（１１０４）に出力し、必要なら画像表示手段（１１０６）にも出力する。

【００５７】再構成参照画像蓄積部（１１０５）は、再構成参照画像を蓄積し、予測画像復号化部（１１０５）に出力する。予測画像復号化部（１１０５）では、予測画像符号化データを復号化し、再構成予測画像を画像表示部（１００６）に出力する。画像表示部（１１０６）は入力された画像を蓄積しておき、要求により表示する。

【００５８】以上のように構成された本実施の形態の画像復号化装置の動作を説明する。データ記憶・伝送部（１１０１）では、参照関係テーブルと符号化データを蓄積・伝送し、参照テーブルと符号化データを符号化データ選択部（１１０２）に出力する。

【００５９】符号化データ選択部（１１０２）では、参照関係テーブルから符号化データ位置を検索し、データ記憶・伝送部（１１０１）から、符号化データ位置の符号化データを取得する。参照／予測フラグと参照符号化データを、参照画像復号化部（１１０３）に出力し、予測符号化データがあれば、予測画像復号化部（１１０５）に出力する。但し、参照画像復号化部（１１０３）に出力する参照符号化データが前回と同じ場合は、参照画像はすでに復号化され、蓄積されているので、参照画像復号化部（１１０３）への出力は省略する。参照画像復号化部（１１０３）では、参照画像が符号化された参照符号化データを復号化し、再構成参照画像を再構成参照画像蓄積部（１１０４）に出力する。

【００６０】また、参照／予測フラグにおいて、参照画像であれば、画像表示部（１００６）に再構成参照画像を出力する。再構成参照画像蓄積部（１１０４）では、再構成された参照画像を蓄積し、予測画像復号化部（１１０５）に出力する。

【００６１】予測画像復号化部（１１０５）では、予測画像が符号化された予測符号化データを再構成参照画像を用いて復号化し、画像表示部（１１０６）に再構成予測画像を出力する。

【００６２】画像表示部（１１０６）では、復号化された画像を蓄積しておき、計算機に接続されたマウスにより検索し、表示する。

【００６３】以上のように、本実施の形態では、多くの異なる視点から撮影した画像を、撮影位置が近傍の画像をグループ化し、グループを予測符号化された符号化データを、その参照予測関係と符号化データ位置と撮影位置を記述したテーブルを用いることにより、効率的に復号化することが可能である。

【００６４】また、参照関係テーブルはグループ順に記述されているので、同一参照予測グループ内が連続的に復号化されるので、再構成参照画像を蓄積しておき、予測符号化データを復号化するだけで、効率的に画像を復号化することが可能である。

【００６５】図１２は本発明の第３の実施の形態である画像復号化装置の構成を示すブロック図であり、同図を用いて本実施の形態の構成を説明する。

【００６６】同図において、データ記憶・伝送部（１２０１）では、参照関係テーブルと符号化データを蓄積・伝送し、参照テーブルをランダムアクセステーブル作成部（１２０２）に出力し、符号化データを符号化データ選択部（１２０３）に出力する。画像要求部（１２０４）では、要求する画像の撮影位置を符号化データ選択部（１２０３）に出力する。ランダムアクセステーブル作成部（１２０２）では、撮影位置から符号化データに効率的にランダムアクセスできるように、テーブルを変更する。

【００６７】符号化データ選択部（１２０３）では、画像要求に基づいて、ランダムアクセステーブルから、必要なデータの位置を検索し、符号化データから必要なデータを参照画像復号化部（１２０５）や予測画像復号化部（１２０７）に出力する。参照画像復号化部（１２０５）では、参照画像符号化データを復号化し、再構成参照画像を再構成参照画像蓄積部（１２０６）に出力し、必要なら画像表示部（１２０８）にも出力する。

【００６８】再構成参照画像蓄積部（１２０６）は、再構成参照画像を蓄積し、予測画像復号化部（１２０７）に出力する。予測画像復号化部（１２０７）では、予測画像符号化データを復号化し、再構成予測画像を画像表示部（１２０７）に出力する。画像表示部（１２０８）は入力された画像を表示する。

【００６９】以上のように構成された本実施の形態の画像復号化装置の動作を説明する。データ記憶・伝送部（１２０１）では、参照関係テーブルと符号化データを蓄積・伝送し、参照テーブルをランダムアクセステーブル作成部（１２０２）に出力する。また、符号化データ選択部（１２０３）の要求により、符号化データを符号化データ選択部（１２０３）に出力する。

【００７０】画像要求部（１２０４）では、計算機に接続されたマウスを操作することによって、ユーザが見たい視点位置を指定し、その視点位置に対応する撮影位置

を符号化データ選択部（１２０３）に出力する。

【００７１】ランダムアクセステーブル作成部（１２０２）では、図９に示す参照関係テーブルを、撮影位置から符号化データ位置を効率的に検索できるよう、ランダムアクセステーブルを作成する。図９中の参照ヘッダと予測ヘッダは、図１３のようなランダムアクセスヘッダ（Ｎバイト）に変換され、撮影位置（ρ，θ）のランダムアクセスヘッダはランダムアクセステーブルの先頭から、

【００７２】

【数６】

$$(359\rho+\theta)N$$

【００７３】の位置に配置し、図１４に示すランダムアクセステーブルを作成する。ただし、参照画像の場合、予測符号化データ位置の欄は空欄にしておく。

【００７４】符号化データ選択部（１２０３）では、要求画像の撮影位置データ（α，β）を用いて、ランダムアクセステーブルの先頭から、

【００７５】

【数７】

$$(359\alpha+\beta)N$$

【００７６】の位置のランダムアクセスヘッダを検索し、データ記憶・伝送手段（１２０１）から、符号化データ位置の符号化データを取得する。参照／予測フラグと参照符号化データは、参照画像復号化部（１２０５）に出力し、予測符号化データがあれば、予測画像復号化部（１２０７）に出力する。

【００７７】ただし、参照画像復号化部（１２０５）に出力する参照符号化データが前回と同じ場合は、参照画像はすでに復号化され、蓄積されているので、参照画像復号化部（１２０５）への出力は省略する。

【００７８】参照画像復号化部（１２０５）では、参照画像が符号化された参照符号化データを復号化し、再構成参照画像を再構成参照画像蓄積部（１２０６）に出力する。また、参照／予測フラグにおいて、参照画像であれば、画像表示部（１２０８）に再構成参照画像を出力する。

【００７９】再構成参照画像蓄積部（１２０６）では、再構成された参照画像を蓄積し、予測画像復号化部（１２０７）に出力する。

【００８０】予測画像復号化部（１１０７）では、予測画像が符号化された予測符号化データを再構成参照画像を用いて復号化し、画像表示部（１２０８）に再構成予測画像を出力する。画像表示部（１２０８）では、入力された画像を表示する。

【００８１】以上のように、本実施の形態では、多くの異なる視点から撮影した画像を、撮影位置が近傍の画像をグループ化し、効率的に予測符号化し、その参照予測関係と符号化データ位置と撮影位置を記述したテーブルを付加することにより、任意の撮影位置の画像を効率的に検索、復号化、表示することが可能である。また、連続的に要求される画像は、撮影位置が近傍のものが多く、同一参照予測グループ内である場合が多く、再構成参照画像を蓄積しておくことで、予測符号化データを復号化するだけで、効率的に画像を復号化、表示することが可能である。

【００８２】図１５は本発明の第４の実施の形態である画像符号化装置の構成を示すブロック図であり、同図を用いて本実施の形態の構成を説明する。

【００８３】同図において、撮影位置制御型画像撮像部（１５０１）は、対象物体の画像を撮影し、その撮影位置データと画像データを出力する。参照予測グループ部（１５０２）は、撮影位置データと画像データを画像データを参照画像と一枚以上の予測画像からなる参照予測グループにグループ化し、参照画像を参照画像符号化部（１５０３）に出力し、予測画像を予測画像符号化部（１５０５）に出力する。参照画像符号化部（１５０３）は、参照画像を符号化し、再構成参照画像を作成し、符号化データを参照関係テーブル作成部（１５０６）に出力し、再構成参照画像を再構成参照画像蓄積部（１５０４）に出力する。再構成参照画像蓄積部（１５０４）は、再構成参照画像を蓄積し、予測画像符号化部（１５０５）に再構成参照画像を出力する。予測画像符号化部（１５０５）は、予測画像を符号化し、符号化データを参照関係テーブル作成部（１５０６）に出力する。参照関係テーブル作成部（１５０６）は、符号化データから参照関係と符号化データ位置と撮影位置を記述した参照関係テーブルを作成し、参照関係テーブルと符号化データをランダムアクセステーブル作成部（１５０７）に出力する。

【００８４】ランダムアクセステーブル作成部（１５０７）では、撮影位置から符号化データに効率的にランダムアクセスできるように、テーブルを変更し、ランダムアクセステーブルと符号化データをデータ記録・伝送部（１５０８）に出力する。データ記録・伝送部（１５０８）は、用途に応じて記録・伝送する。

【００８５】以上のように構成された本実施の形態の画像符号化装置の動作を説明する。まず、画像位置制御型画像撮像部（１５０１）は、画像位置制御型画像撮影部（１０１）と同様である。参照予測グループ化部（１５０２）は、参照予測グループ化部（１０５）と同様である。参照画像符号化部（１５０３）は、参照画像符号化部（１０６）と同様である。再構成参照画像蓄積部（１５０４）は、再構成参照画像蓄積部（１０７）と同様である。予測画像符号化部（１５０５）は、予測画像符号化部（１０８）と同様である。参照関係テーブル作成部（１５０６）は、参照関係テーブル作成部（１０９）と同様である。

【００８６】ランダムアクセステーブル作成部（１５０７）は、ランダムアクセステーブル作成部（１２０２）

と同様に、ランダムアクセステーブルを作成し、データ記録・伝送部（１１０）に符号化データとランダムアクセステーブルを出力する。

【００８７】データ記録・伝送部（１１０）では、ランダムアクセステーブルと符号化データを磁気ディスクや光ディスクに蓄積しておいたり、ネットワークなどで伝送する。

【００８８】以上のように、本実施の形態では、異なる多くの視点から撮影された多視点画像を、撮影位置データが近傍の類似した画像である参照画像と１枚以上の予測画像にグループ化し、効率よく符号化することが可能である。さらに、画像を各参照予測グループで参照画像が先頭になるように整列して、符号化手段に出力することで、計算量と計算に必要な蓄積容量を少なくしている。また、その撮影位置順に符号化データ位置を格納したテーブルを付加することで、ランダムアクセスに対応した効率的な復号化が可能である。

【００８９】図１６は本発明の第５の実施の形態である画像復号化装置の構成を示すブロック図であり、同図を用いて本実施の形態の構成を説明する。

【００９０】同図において、データ記憶・伝送部（１６０１）では、ランダムアクセステーブルと符号化データを蓄積・伝送し、ランダムアクセステーブルと符号化データを符号化データ選択部（１６０２）に出力する。画像要求部（１６０３）では、要求する画像の撮影位置を符号化データ選択部（１６０２）に出力する。

【００９１】符号化データ選択部（１６０２）では、画像要求に基づいて、ランダムアクセステーブルから、必要なデータの位置を検索し、符号化データから必要なデータを参照画像復号化部（１６０４）や予測画像復号化部（１６０６）に出力する。参照画像復号化部（１６０４）では、参照画像符号化データを復号化し、再構成参照画像を再構成参照画像蓄積部（１６０５）に出力し、必要なら画像表示部（１６０７）にも出力する。

【００９２】再構成参照画像蓄積部（１６０５）は、再構成参照画像を蓄積し、予測画像復号化部（１６０６）に出力する。予測画像復号化部（１６０６）では、予測画像符号化データを復号化し、再構成予測画像を画像表示部（１６０７）に出力する。画像表示部（１６０７）は入力された画像を表示する。

【００９３】以上のように構成された本実施の形態の画像復号化装置の動作を説明する。データ記憶・伝送部（１６０１）では、ランダムアクセステーブルと符号化データを蓄積・伝送し、ランダムアクセステーブルを符号化データ選択部（１６０２）に出力する。また、符号化データ選択部（１６０２）の要求により、符号化データを符号化データ選択部（１６０２）に出力する。

【００９４】画像要求部（１６０３）では、計算機に接続されたマウスを操作することによって、ユーザが見たい視点位置を指定し、その視点位置に対応する撮影位置

を符号化データ選択部（１６０２）に出力する。符号化データ選択部（１６０２）は、符号化データ選択部（１２０３）と同様である。

【００９５】参照画像復号化部（１６０４）では、参照画像が符号化された参照符号化データを復号化し、再構成参照画像を再構成参照画像蓄積部（１６０５）に出力する。また、参照／予測フラグにおいて、参照画像であれば、画像表示部（１６０７）に再構成参照画像を出力する。再構成参照画像蓄積部（１６０５）では、再構成された参照画像を蓄積し、予測画像復号化部（１６０６）に出力する。

【００９６】予測画像復号化部（１６０６）では、予測画像が符号化された予測符号化データを再構成参照画像を用いて復号化し、画像表示部（１６０７）に再構成予測画像を出力する。画像表示部（１６０７）では、入力された画像を表示する。

【００９７】以上のように、本実施の形態では、多くの異なる視点から撮影した画像を、撮影位置が近傍の画像をグループ化し、効率的に予測符号化し、その参照予測関係と符号化データ位置と撮影位置を記述したテーブルを付加することにより、任意の撮影位置の画像を効率的に検索、復号化、表示することが可能である。また、連続的に要求される画像は、撮影位置が近傍のものが多く、同一参照予測グループ内である場合が多く、再構成参照画像を蓄積しておくことで、予測符号化データを復号化するだけで、効率的に画像を復号化、表示することが可能である。

【００９８】図１７は本発明の第６の実施の形態である画像復号化装置の構成を示すブロック図であり、同図を用いて本実施の形態の構成を説明する。

【００９９】同図において、データ記憶・伝送部（１７０１）では、ランダムアクセステーブルと符号化データを蓄積・伝送し、ランダムアクセステーブルと符号化データを符号化データ選択部（１７０５）に出力する。画像要求監視部（１７０２）は、画像の要求を監視し画像要求蓄積部（１７０３）に出力する。

【０１００】画像要求蓄積部（１７０３）は、画像要求を蓄積する。復号化状況蓄積部（１７０４）は復号化の状況を蓄積する。符号化データ選択部（１７０５）は、復号化状況と画像要求とランダムアクセステーブルと符号化データから復号化すべき符号化データを選択し、参照画像復号化部（１７０６）や変形画像復号化部（１７０８）や予測画像復号化部（１７１０）や出力画像選択部（１７１２）に出力する。参照画像復号化部（１７０６）は、参照画像符号化データを復号化し、再構成参照画像を再構成参照画像蓄積部（１７０７）に出力する。

【０１０１】再構成参照画像蓄積部（１７０７）は再構成参照画像を蓄積する。変形参照画像復号化部（１７０８）は符号化データの動きベクトル部分と再構成参照画像から、変形参照画像を復号化し、変形参照画像蓄積部

（１７０９）に出力する。

【０１０２】変形参照画像蓄積部（１７０９）は、変形参照画像を蓄積する。予測画像復号化部（１７１０）は、符号化データの残差成分部分と変形参照画像から、予測画像を復号化し、予測画像蓄積部（１７１１）に出力する。予測画像蓄積部（１７１１）は、予測画像を蓄積する。出力画像選択部（１７１２）は、画像を選択し、画像表示部（１７１３）に出力する。画像表示部（１７１３）は、画像を表示する。

【０１０３】以上のように構成された本実施の形態の画像復号化装置の動作を説明する。まず、データ記憶・伝送部（１７０１）では、ランダムアクセステーブルと符号化データを蓄積・伝送し、ランダムアクセステーブルを符号化データ選択部（１７０５）に出力する。また、符号化データ選択部（１７０５）により、符号化データが読み出される。

【０１０４】画像要求監視部（１７０２）は、マウスを持った計算機であり、ユーザがマウスを使ってする要求を常に監視し、画像要求蓄積部（１７０３）に蓄積する。画像要求蓄積部（１７０３）は、要求される画像の撮影位置を蓄積し、符号化データ選択部（１７０５）により読み出される。

【０１０５】復号化状況蓄積部（１７０４）では、復号化状況テーブル（１８０１）を蓄積し、符号化データ選択部（１７０４）により更新され、読み出される。復号化状況テーブル（１８０１）において、撮影位置は（ρ，θ）であり、復号化レベルは復号化の状況を示す。復号化レベルの意味を復号化レベルテーブル（１８０２）に示す。

【０１０６】符号化データ選択部（１７０５）は、要求する撮影位置とランダムアクセステーブルと復号化状況テーブルから、出力データと出力先と復号化状況テーブル更新方法を決定する。図１９に要求する撮影位置とランダムアクセステーブルと復号化状況テーブルの状態を列挙した状態テーブルを示す。ここで、Ｐは要求された画像の撮影位置、Ｑは要求された画像が予測画像の場合、その画像の復号化に必要な参照画像の撮影位置、Ｒは復号化状況テーブル（１８０１）に登録されている参照画像撮影位置、Ｓは復号化状況テーブル（１８０１）に登録されている予測画像撮影位置、Ｌは復号化状況テーブル（１８０１）に登録されている復号化レベルを表わす。

【０１０７】図２０に出力データ書式（２００１）と各状態に対応する出力データと出力先と更新後の復号化状況テーブルを列挙した出力データテーブル（２００２）を示す。出力データ書式（２００１）において出力画像フラグは最終的に出力されるべき画像を示すフラグであり、０は出力なし、１は再構成参照画像を出力、２は変形参照画像を出力、３は再構成予測画像を出力を示す。符号化データはランダムアクセステーブルより撮影位置に対応する符号化データ位置を検索し、データ記憶・伝送部（１７０１）より読み出す。

【０１０８】参照画像復号化部（１７０６）では、入力された出力画像フラグと符号化データを出力画像フラグは出力画像選択部（１７１２）に出力し、符号化データは参照画像に復号化し、再構成参照画像蓄積部（１７０７）に出力する。再構成参照画像蓄積部（１７０７）は、再構成参照画像を蓄積する。変形画像復号化部（１７０８）では、入力された出力画像フラグと符号化データを出力画像フラグは出力画像選択部（１７１２）に出力し、符号化データは符号化データの動きベクトル部分を復号化し、再構成参照画像を変形させて変形参照画像を作成し、変形参照画像蓄積部（１７０９）に出力する。変形参照画像蓄積部（１７０９）は、変形参照画像を蓄積する。

【０１０９】予測画像復号化部（１７１０）では、入力された出力画像フラグと符号化データを、出力画像フラグは出力画像選択部（１７１２）に出力し、符号化データは符号化データの残差成分部分を復号化し、変形参照画像に残差成分を加えて予測画像を作成し、再構成予測画像蓄積部（１７１１）に出力する。再構成予測画像蓄積部（１７１１）は、再構成予測画像を蓄積する。

【０１１０】出力画像選択部（１７１２）は、入力された出力画像フラグが、０ならば出力しない、１ならば再構成参照画像蓄積部（１７０６）より再構成参照画像を読み出し、画像表示部（１７１３）に出力する、２ならば変形参照画像蓄積部（１７０８）より変形参照画像を読み出し、画像表示部（１７１３）に出力する、３ならば再構成予測画像蓄積部（１７１１）より再構成予測画像を読み出し、画像表示部（１７１３）に出力する。画像表示部（１７１３）は、入力された画像を表示する。

【０１１１】以上のように、本実施の形態では、画像要求を監視し、蓄積する部と、復号化処理途中の変形画像を出力する部を備えており、画像要求の変化の速度と画像復号化処理の速度に対応して、自動的に変形画像か再構成予測画像を出力することにより、画像要求の変化が速いときや画像復号化処理の速度が遅いときは画質の良くない変形画像が、その逆のときには画質のよい再構成予測画像を表示することが可能である。

【０１１２】図２１は本発明の第７の実施の形態である画像復号化装置の構成を示すブロック図であり、同図を用いて本実施の形態の構成を説明する。

【０１１３】同図において、データ記憶・伝送部（２１０１）では、ランダムアクセステーブルと符号化データを蓄積・伝送し、ランダムアクセステーブルと符号化データを符号化データ選択部（２１０４）に出力する。画像要求部（２１０２）は、要求する画像の撮影位置を画像蓄積検査部（２１０３）に出力する。画像蓄積検査部（２１０３）では、再構成参照画像群蓄積部（２１０８）と再構成予測画像群蓄積部（２１０９）に蓄積され

ている画像の重要度を撮影位置に基づいて計算し、最大蓄積量を越えないように重要度の低い画像を消去する。符号化データ選択部（２１０４）では、蓄積された画像とランダムアクセステーブルに基づいて必要な符号化データを選択し、参照画像復号化部（２１０５）や予測画像復号化部（２１０６）や参照予測画像復号化部（２１０７）に出力する。

【０１１４】再構成参照画像群蓄積部（２１０８）では、複数枚の再構成参照画像と各画像の撮影位置を蓄積し、出力画像選択部（２１１０）に出力する。再構成予測画像群蓄積部（２１０９）では、複数枚の再構成予測画像と各画像の撮影位置を蓄積し、出力画像選択部（２１１０）に出力する。参照画像復号化部（２１０５）では、符号化データを参照画像に復号化し、再構成参照画像群蓄積部（２１０８）に出力し、要求画像の撮影位置を出力画像選択部（２１１０）に出力する。予測画像復号化部（２１０６）では、符号化データを再構成参照画像を用いて予測画像に復号化し、再構成予測画像群蓄積部（２１０９）に出力し、要求画像の撮影位置を出力画像選択部（２１１０）に出力する。参照予測画像復号化部（２１０７）では、符号化データを参照画像と予測画像に復号化し、再構成した参照画像を再構成参照画像群蓄積部（２１０８）に出力し、再構成した予測画像を再構成予測画像群蓄積部（２１０９）に出力し、要求画像の撮影位置を出力画像選択部（２１１０）に出力する。出力画像選択部（２１１０）は、要求画像の撮影位置に対応する画像を再構成参照画像群蓄積部（２１０８）と再構成予測画像群蓄積部（２１０９）の中から取得し、画像表示部（２１１１）に出力する。画像表示部（２１１１）では、入力された画像を表示する。

【０１１５】以上のように構成された本実施の形態の画像復号化装置の動作を説明する。データ記憶・伝送部＊

＊（２１０１）では、ランダムアクセステーブルと符号化データを蓄積・伝送し、ランダムアクセステーブルを符号化データ選択部（２１０４）に出力する。また、符号化データ選択部（２１０４）により、符号化データが読み出される。

【０１１６】画像要求部（２１０２）は、マウスを持った計算機であり、ユーザがマウスを使って見たい画像を要求し、要求画像の撮影位置を画像蓄積検査手段（２１０３）に出力する。画像蓄積検査部（２１０３）では、要求された画像の撮影位置と、最大Ｎ枚蓄積できる再構成参照画像群蓄積部（２１０８）に蓄積された画像の撮影位置と最大Ｍ枚蓄積できる再構成予測画像群蓄積部（２１０９）に蓄積された画像の撮影位置により各画像に対する重要度を計算し、再構成参照画像群蓄積部（２１０８）に蓄積された画像のＮ番目に重要度が大きいものと再構成予測画像群蓄積部（２１０９）に蓄積された画像のＭ番目に重要度が大きいものを消去する。$\rho-\theta$空間上で距離の近いほど重要度は高いので、要求された画像の撮影位置をＰ、蓄積された画像の撮影位置をＱとすると、重要度はＷは、

【０１１７】

【数８】

$$W = \frac{1}{d(\mathbf{P},\mathbf{Q})}$$

【０１１８】で計算できる。ここで、

【０１１９】

【数９】

$$\mathbf{P}_1 = (\rho_1, \theta_1),\ \mathbf{P}_2 = (\rho_2, \theta_2)$$

【０１２０】とすると、

【０１２１】

【数１０】

$$d(\mathbf{P}_1, \mathbf{P}_2) = \min((\rho_1 - \rho_2)^2, (\rho_1 - \rho_2 - 360)^2, (\rho_1 - \rho_2 + 360)^2) + (\theta_1 - \theta_2)^2$$

【０１２２】である。符号化データ選択部（２１０４）では、図２２の状態テーブルに応じて図２３の出力データテーブルにしたがって、出力データを作成し、それぞれの出力先に出力する。図２２と図２３においてＰは要求画像の撮影位置であり、Ｑは要求画像が予測画像の場合に復号化に必要となる参照画像の撮影位置である。出力データは先頭に出力画像撮影位置データを配置し、続いて画像の再構成に必要なデータを配置している。

【０１２３】参照画像復号化部（２１０５）では、入力されたデータの先頭の出力画像撮影位置データを出力画像選択部（２１１０）に出力し、参照符号化データを参照画像に復号化し、再構成参照画像群蓄積部（２１０８）に出力する。

【０１２４】予測画像復号化部（２１０６）では、入力されたデータの先頭の出力画像撮影位置データを出力画像選択部（２１１０）に出力し、参照画像データを用いて予測符号化データを予測画像に復号化し、再構成予測画像群蓄積部（２１０９）に出力する。

【０１２５】参照予測画像復号化部（２１０７）では、入力されたデータの先頭の出力画像撮影位置データを出力画像選択部（２１１０）に出力し、参照符号化データを参照画像に復号化し、再構成された参照画像を用いて予測符号化データを予測画像に復号化し、再構成参照画像群蓄積部（２１０８）に再構成参照画像を出力し、再構成予測画像群蓄積部（２１０９）に再構成予測画像を出力する。

【０１２６】出力画像選択部（２１１０）では、入力された出力画像撮影位置データに対応する画像を再構成参照画像群蓄積部（２１０８）と再構成予測画像群蓄積部（２１０９）から取得し、画像表示部（２１１１）に出力する。

【０１２７】画像表示部（２１１１）では、入力された画像を表示する。以上のように、本実施の形態では、多視点画像復号化装置では、現在要求されている画像の撮

影位置の近傍の画像の要求頻度が高いことを利用し、要求された画像の撮影位置に近いものに高い重要度を与え、一度復号化した画像のうち重要度の高いものを蓄積しておくことで、効率的に画像を復号化することが可能である。

【０１２８】図２４は本発明の第８の実施の形態である画像復号化装置の構成を示すブロック図であり、同図を用いて本実施の形態の構成を説明する。

【０１２９】同図において、データ記憶・伝送部（２４０１）では、ランダムアクセステーブルと符号化データを蓄積・伝送し、ランダムアクセステーブルと符号化データを符号化データ選択部（２４０５）と予測符号化データ選択部（２４０９）に出力する。画像要求部（２４０２）は、要求する画像の撮影位置を画像要求蓄積部（２４０３）に出力する。画像要求蓄積部（２４０３）は、画像要求を蓄積し、符号化データ選択部（２４０５）により画像要求が読み出される。前回画像要求蓄積部（２４０４）は、前回処理した画像の撮影位置を蓄積し、符号化データ選択部（２４０５）によって、前回画像要求が読み出され、更新される。

【０１３０】符号化データ選択部（２４０５）は、前回処理した画像の要求や画像の蓄積の状態から復号化すべき符号化データを決定し、ランダムアクセステーブルと符号化データから必要な符号化データや画像データを選択し、参照予測画像復号化部１（２４０６）や参照画像復号化部１（２４０７）や予測画像復号化部１（２４０８）や画像表示部（２４１５）に出力する。

【０１３１】予測符号化データ選択部（２４０９）は、復号化されていない近傍の撮影位置を予測し、その位置の画像の復号化に必要なデータを、ランダムアクセステーブルと符号化データと画像データからを選択し、参照予測画像復号化部２（２４１０）や参照画像復号化部２（２４１１）や予測画像復号化部２（２４１２）に出力する。参照予測画像復号化部１（２４０６）は、符号化データから予測画像を復号化し、画像表示部（２４１５）に出力する。

【０１３２】参照画像復号化部１（２４０７）は、符号化データから参照画像を復号化し、画像表示部（２４１５）に出力する。予測画像復号化部１（２４０８）は、符号化データと画像データから予測画像を復号化し、画像表示部（２４１５）に出力する。参照予測画像復号化部２（２４１０）は、符号化データから参照画像と予測画像を復号化し、再構成参照画像群蓄積部（２４１３）と再構成予測画像群蓄積部（２４１４）に出力する。参照画像復号化部２（２４１１）は、符号化データから参照画像を復号化し、再構成参照画像群蓄積部（２４１３）に出力する。予測画像復号化部２（２４１２）は、符号化データと画像データから予測画像を復号化し、再構成予測画像群蓄積部（２４１４）に出力する。

【０１３３】以上のように構成された本実施の形態の画像復号化装置の動作を説明する。データ記憶・伝送部（２４０１）では、ランダムアクセステーブルと符号化データを蓄積・伝送し、ランダムアクセステーブルを符号化データ選択部（２４０５）に出力する。また、符号化データ選択部（２４０５）により、符号化データを読み出される。画像要求部（２４０２）は、マウスを持った計算機であり、ユーザがマウスを使って見たい画像を要求し、要求画像の撮影位置を画像要求蓄積手段（２４０３）に出力する。画像要求蓄積部（２４０３）では、要求された画像の撮影位置を蓄積し、符号化データ選択部（２４０５）によって、撮影位置を読み出される。

【０１３４】前回要求画像蓄積部（２４０４）では、処理した撮影位置を蓄積し、符号化データ選択部（２４０５）によって、撮影位置を読み出される。符号化データ選択部（２４０５）では、図２５の状態テーブルに応じて図２６の出力データテーブルにしたがって、出力データを作成し、それぞれの出力先へ出力する。図２５と図２６においてＰは要求画像の撮影位置であり、Ｑは要求画像が予測画像の場合に復号化に必要となる参照画像の撮影位置であり、Ｒは前回処理された画像の撮影位置である。さらに、Ｒ≠Ｐならば、再構成予測画像群蓄積部（２４１３）と再構成参照画像群蓄積部（２４１４）に蓄積されている画像を消去し、Ｐを前回要求画像蓄積部（２４０４）に書き込む。

【０１３５】参照予測画像復号化部１（２４０６）は、参照符号化データと予測符号化データを予測画像に復号化し、画像表示部（２４１５）に出力する。参照画像復号化部１（２４０７）は、参照符号化データを参照画像に復号化し、画像表示部（２４１５）に出力する。予測画像復号化部１（２４０８）は、参照画像データと予測符号化データを予測画像に復号化し、画像表示部（２４１５）に出力する。

【０１３６】予測符号化データ選択部（２４０９）では、図２７の予測テーブルの上から順に、再構成参照画像群蓄積部（２４１３）と再構成予測画像群蓄積部（２４１４）に蓄積されていないか検索し、蓄積されていないものがあれば、その撮影位置を予測撮影位置としてＳとして、図２８の状態テーブルに応じて図２９の出力データテーブルにしたがって、出力データを作成しそれぞれの出力先へ出力する。図２８と図２９で、Ｔは、撮影位置Ｓの画像が予測画像の場合の復号化に必要な参照画像の撮影位置である。

【０１３７】参照予測画像復号化部２（２４１０）は、参照符号化データと予測符号化データを参照画像と予測画像に復号化し、参照画像は再構成参照画像群蓄積部（２４１３）に、予測画像は再構成予測画像群蓄積部（２４１４）に出力する。参照画像復号化部２（２４１１）は、参照符号化データを参照画像に復号化し、再構成参照画像群蓄積部（２４１３）に出力する。

【０１３８】予測画像復号化部２（２４１２）は、参照

画像データと予測符号化データを予測画像に復号化し、再構成予測画像群蓄積部（２４１４）に出力する。画像表示部（２４１５）は、入力された画像を表示する。

【０１３９】以上のように、本実施の形態では、多視点画像復号化装置では、現在要求されている画像の撮影位置の近傍の画像の要求頻度が高いことを利用し、画像の要求の変化がないときには、要求された画像の撮影位置に近いもの復号化しておくことで、効率的に画像を復号化することが可能である。

【０１４０】図３０は本発明の第９の実施の形態である画像符号化装置の構成を示すブロック図であり、同図を用いて本実施の形態の構成を説明する。

【０１４１】同図において、撮影位置制御型画像撮像部（３００１）は、対象物体の画像を撮影し、その撮影位置データと画像データを出力する。参照予測グループ部（３００５）は、撮影位置データと画像データを参照画像と一枚以上の予測画像からなる参照予測グループにグループ化し、参照画像を参照画像符号化部（３００６）に出力し、撮影位置データと参照画像と予測画像を変形拘束式推定部（３００８）に出力する。

【０１４２】参照画像符号化部（３００６）は、参照画像を符号化し、再構成参照画像を作成し、符号化データを参照関係作成部（３０１０）に出力し、再構成参照画像を再構成参照画像蓄積部（３００７）に出力する。再構成参照画像蓄積部（３００７）は、再構成参照画像を蓄積し、拘束予測画像符号化部（３００９）に再構成参照画像を出力する。拘束式推定部（３００８）は、拘束式を推定し、拘束予測画像符号化部（３００９）に出力する。

【０１４３】拘束予測画像符号化部（３００９）は、予測画像を符号化し、符号化データを参照関係作成部（３０１０）に出力する。参照関係テーブル作成部（３０１０）は、符号化データから参照関係と符号化データ位置と撮影位置を記述した参照関係テーブルを作成し、参照関係テーブルと符号化データをデータ記録・伝送部（３０１１）に出力する。データ記録・伝送部（３０１１）は、用途に応じて記録・伝送する。

【０１４４】以上のように構成された本実施の形態の画像符号化装置の動作を説明する。まず、画像位置制御型撮像部（３００１）では、図１の画像位置制御型撮像部と同様に、対象物体の周囲を周囲を網羅するように、異なる視点から画像を撮影し、撮影位置・画像データを取得し、図３に示す書式で出力する。

【０１４５】参照グループ化部（３００５）では、図１の参照グループ化部（１０５）と同様に、参照画像とその近傍の画像を予測画像としてグループ化し、インデックス順に図５に示す書式で、参照画像符号化部（３００６）と変形拘束式推定部（３００８）へ出力する。参照画像符号化部（３００６）は、図１の参照画像符号化部（１０６）と同様である。再構成参照画像蓄積部（３００７）は、図１の再構成参照画像蓄積部（１０７）と同様である。

【０１４６】変形拘束式推定部（３００８）では、エピポーラ拘束と呼ばれる拘束式のパラメータ７つを推定する（「エピポラー　ジェオメトリ　イン　ステレオ　モーション　アンド　オブジェクト　レコグニション」"Epipolar Geometry in Stereo、 Motion and Object Recognition - A Unified Approach、" Gang Xu and Zheng you Zhang、 Kluwer Academic Publishers、 1996））。

【０１４７】図３１はエピポーラ拘束の説明図であり、３次元空間中の点Ｍを異なる視点から撮影した場合である。Ｃ１、Ｃ２はそれぞれの画像１と画像２を撮影したカメラの焦点を示し、Ｍは、画像１上ではｍ１に、画像２上ではｍ２に投影される。Ｃ２は画像１上のｅ１に、Ｃ１は画像２上のｅ２に投影され、ｅ１とｅ２はそれぞれの画像上のエピ極と呼ばれる。３次元空間中のＣ１とｍ１を通る直線を画像２上への投影した直線をエピポーラ線と呼び、エピポーラ線２として図３１に示す。

【０１４８】エピポーラ線２は、画像１を撮影したカメラの内部パラメータ（焦点距離、光軸）と視点位置、画像２を撮影したカメラの内部パラメータと視点位置、画像１上の点の座標により決定できる。エピポーラ線２を決定できれば、画像１上の点ｍ１に３次元空間で対応する画像２上の点ｍ２の位置は、Ｃ１からＭの距離が得られない場合でも、エピポーラ線２上に拘束できる。

【０１４９】図３２に示す二つの視点から撮影された画像１と画像２において、画像１上の点（ｘ１、ｙ１）に対応する画像２上のエピポーラ線は以下のように表わされる。

【０１５０】

【数１１】

$$(x_1, y_1, 1)\mathbf{F}(x, y, 1)^{\mathbf{T}} = 0$$

【０１５１】ここで、

【０１５２】

【数１２】

$$\mathbf{F} = \mathbf{A}^{-T}\mathbf{t}_{\times}\mathbf{R}\mathbf{A}^{-1}$$

【０１５３】

【数１３】

$$\mathbf{A} = \begin{pmatrix} f & 0 & ox \\ 0 & f & oy \\ 0 & 0 & 1 \end{pmatrix}$$

【０１５４】

【数１４】

25 26

$$R=\begin{pmatrix}1 & 0 & 0\\ 0 & \cos(\rho_2-\rho_1) & -\sin(\rho_2-\rho_1)\\ 0 & \sin(\rho_2-\rho_1) & \cos(\rho_2-\rho_1)\end{pmatrix}\begin{pmatrix}\cos\theta_2 & 0 & -\sin\theta_2\\ 0 & 1 & 0\\ \sin\theta_2 & 0 & \cos\theta_2\end{pmatrix}\begin{pmatrix}\cos\theta_1 & 0 & \sin\theta_1\\ 0 & 1 & 0\\ -\sin\theta_1 & 0 & \cos\theta_1\end{pmatrix}$$

【０１５５】

【数１５】

$$\mathbf{t}=\mathbf{R}(0,0,-d)^T+(0,0,d)^T$$

【０１５６】であり、ｏｘ、ｏｙはカメラの光軸の画像上の座標を表わし、ｆは焦点距離を示す。

【０１５７】（$\rho 1$，$\theta 1$）、（$\rho 2$，$\theta 2$）、ｄは、図３２に示す撮影位置を示し、×は、

【０１５８】

【数１６】

$$\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}_\times=\begin{pmatrix}0 & -x_3 & x_2\\ x_3 & 0 & -x_1\\ -x_2 & x_1 & 0\end{pmatrix}$$

【０１５９】の演算を示す。カメラの内部パラメータであるｏｘ、ｏｙ、ｆは、予め求めておき、（$\rho 1$，$\theta 1$）、（$\rho 2$，$\theta 2$）、ｄは撮影システムにより与えられる。７つのパラメータｅは、

【０１６０】

【数１７】

$$\mathbf{F}=\begin{pmatrix}f_{11} & f_{12} & f_{13}\\ f_{21} & f_{22} & f_{23}\\ f_{31} & f_{32} & f_{33}\end{pmatrix}$$

【０１６１】とすると、

【０１６２】

【数１８】

$$\mathbf{e}=\begin{pmatrix}\frac{f_{32}f_{21}-f_{22}f_{31}}{f_{22}f_{11}-f_{21}f_{12}}\\ \frac{f_{31}f_{12}-f_{11}f_{32}}{f_{22}f_{11}-f_{21}f_{12}}\\ \frac{f_{23}f_{12}-f_{22}f_{13}}{f_{22}f_{11}-f_{21}f_{12}}\\ \frac{f_{13}f_{21}-f_{11}f_{23}}{f_{22}f_{11}-f_{21}f_{12}}\\ \frac{-f_{12}}{f_{21}}\\ \frac{f_{11}}{f_{21}}\\ \frac{-f_{22}}{f_{21}}\end{pmatrix}$$

【０１６３】で計算できる。予測画像の撮影位置を（$\rho 1$，$\theta 1$）とし、その予測画像を予測するための参照画像（同じインデックスのついた参照画像）の撮影位置（$\rho 2$，$\theta 2$）として用いて、それぞれの参照予測の組に対してエピポーラパラメータを推定する。入力である図５のデータの予測画像にエピポーラパラメータを付加し、図３３に示す書式で、拘束予測画像符号化部（３００９）へ出力する。

【０１６４】拘束予測画像符号化部（３００９）のフローチャートを図３４に示す。ブロック化ステップでは、予測画像データをブロックに分割する。

【０１６５】エピポーラ線計算ステップでは、各ブロックの中心点の座標を（ｘ１，ｙ１）とし、拘束式推定部（３００８）で推定されたエピポーラパラメータｅからＦを以下のように計算し、式１１を用いてブロックにエピポーラ線を求める。

【０１６６】

【数１９】

$$\mathbf{F}=\begin{pmatrix}b & a & -(ay_b+bx_b)\\ -1 & -c & (cy_b+bx_b)\\ y_a-bx_a & cy_a-ax_a & -(cy_b-x_b)y_a+(ay_b+bx_b)x_a\end{pmatrix}$$

【０１６７】ここで、

【０１６８】

【数２０】

27

$$e = \begin{pmatrix} x_a \\ y_a \\ x_b \\ y_b \\ a \\ b \\ c \end{pmatrix}$$

【０１６９】である。動き推定ステップでは、図３５に示すように、予測画像のブロックをエピポーラ線に沿って、再構成参照画像上を１次元的に移動させて、予測画像ブロックと再構成参照画像の画素値の差の絶対値の総和が最小になる１次元の移動量を求める。

【０１７０】ブロック残差計算ステップでは、予測画像ブロックと再構成参照画像の画素値の差の絶対値の総和が最小になる位置での、画素値の差を計算する。ＤＣＴステップでは、ブロック残差計算ステップで計算されたブロック残差をデジタルコサイン変換する。量子化ステップでは、ＤＣＴステップの出力結果を所定の量子幅で量子化する。

【０１７１】可変長符号化ステップでは、動き推定ステップで計算された１次元移動量と、量子化ステップの出力結果を可変長符号化し、符号化データとして図３６に示す書式で出力する。

【０１７２】参照関係テーブル作成部（３０１０）は、参照関係テーブル作成部（１０９）と同様である。データ記録・伝送部（３０１１）は、データ記録・伝送部（１１０）と同様である。

【０１７３】以上のように、本実施の形態では、多視点画像のように、動きのないシーンを与えられた視点から撮影した場合、２つの画像の対応点が撮影対象によらず、カメラの視点と内部パラメータから決定できるある直線上に拘束されることを利用して、符号化データに拘束パラメータを付加する一方で、２次元の動き情報を一次元の動き情報に減少させることで、効率の良い符号化が可能である。

【０１７４】図３７は本発明の第１０の実施の形態である画像復号化装置の構成を示すブロック図であり、同図を用いて本実施の形態の構成を説明する。

【０１７５】同図において、データ記憶・伝送部（３７０１）では、参照関係テーブルと符号化データを蓄積・伝送し、参照テーブルと符号化データを符号化データ選択部（３７０２）に出力する。符号化データ選択部（３７０２）では、参照関係テーブルからデータの位置を検索し、符号化データから必要なデータを参照画像復号化部（３７０３）や拘束予測画像復号化部（３７０５）に出力する。参照画像復号化部（１７０３）では、参照画像符号化データを復号化し、再構成参照画像を再構成参照画像蓄積部（３７０４）に出力し、必要なら画像表示手段（３７０６）にも出力する。

28

【０１７６】再構成参照画像蓄積部（３７０５）は、再構成参照画像を蓄積し、拘束予測画像復号化部（３７０５）に出力する。拘束予測画像復号化部（３７０５）では、予測画像符号化データを復号化し、再構成予測画像を画像表示部（３７０６）に出力する。画像表示部（３７０６）は入力された画像を蓄積しておき、要求により表示する。

【０１７７】以上のように構成された本実施の形態の画像復号化装置の動作を説明する。データ記憶・伝送部（３７０１）では、参照関係テーブルと符号化データを蓄積・伝送し、参照テーブルと符号化データを符号化データ選択部（３７０２）に出力する。符号化データ選択部（３７０２）では、参照関係テーブルから符号化データ位置を検索し、データ記憶・伝送部（３７０１）から、符号化データ位置の符号化データを取得する。参照／予測フラグと参照符号化データを、参照画像復号化部（３７０３）に出力し、予測符号化データがあれば、予測画像復号化部（３７０５）に出力する。

【０１７８】ただし、参照画像復号化部（３７０３）に出力する参照符号化データが前回と同じ場合は、参照画像はすでに復号化され、蓄積されているので、参照画像復号化部（３７０３）への出力は省略する。

【０１７９】参照画像復号化部（３７０３）では、参照画像が符号化された参照符号化データを復号化し、再構成参照画像を再構成参照画像蓄積部（３７０４）に出力する。また、参照／予測フラグにおいて、参照画像であれば、画像表示部（３７０６）に再構成参照画像を出力する。

【０１８０】再構成参照画像蓄積部（３７０４）では、再構成された参照画像を蓄積し、拘束予測画像復号化部（３７０５）に出力する。拘束予測画像復号化部（３７０５）のフローチャートを図３８に示す。復号化ステップでは、可変長符号化されたエピポーラパラメータ、動き情報、残差情報を復号化する。

【０１８１】エピポーラ線計算ステップでは、ブロックの中心座標とエピポーラパラメータから、図３４のエピポーラ線計算ステップと同様にエピポーラ線を計算する。動き補償ステップでは、エピポーラ線に沿って、復号化された移動量だけ移動したところからブロックを得、変形画像を構成する。逆量子化ステップでは、残差情報を逆量子化する。逆ＤＣＴステップでは、逆量子化された残差情報を逆ＤＣＴし、画素の残差情報を得る。ブロック再構成ステップでは、逆ＤＣＴステップで得られた残差情報を、動き補償ステップで得られた変形画像に加えて予測画像データを再構成する。

【０１８２】以上のように予測符号化データを再構成参照画像を用いて復号化し、画像表示部（３７０６）に再構成予測画像を出力する。画像表示部（３７０６）では、復号化された画像を蓄積しておき、計算機に接続されたマウスにより検索し、表示する。

29

【０１８３】以上のように、本実施の形態では、多視点画像のように、動きのないシーンを与えられた視点から撮影した場合、２つの画像の対応点が撮影対象によらず、カメラの視点と内部パラメータから決定できるある直線上に拘束されることを利用して、符号化データに拘束パラメータを付加する一方で、２次元の動き情報を一次元の動き情報に減少させることで、効率の良く符号化した符号化データの復号化が可能である。

【０１８４】図３９は本発明の第１１の実施の形態である画像符号化装置の構成を示すブロック図であり、同図を用いて本実施の形態の構成を説明する。

【０１８５】同図において、撮影位置光源位置制御型画像撮像部（３９０１）は、対象物体の画像を撮影し、その撮影位置データと光源位置データと画像データを出力する。参照予測グループ部（３９０７）は、撮影位置データと光源位置データと画像データを参照画像と一枚以上の予測画像からなる参照予測グループにグループ化し、参照画像を参照画像符号化部（３９０８）に出力し、予測画像を予測画像符号化部（３９０８）に出力する。

【０１８６】参照画像符号化部（３９０８）は、参照画像を符号化し、再構成参照画像を作成し、符号化データを参照関係作成部（３９１１）に出力し、再構成参照画像を再構成参照画像蓄積部（３９０９）に出力する。再構成参照画像蓄積部（３９０９）は、再構成参照画像を蓄積し、予測画像符号化部（３９１０）に再構成参照画像を出力する。

【０１８７】予測画像符号化部（３９１０）は、予測画像を符号化し、符号化データを参照関係作成部（３９１１）に出力する。参照関係テーブル作成部（３９１１）は、符号化データから参照関係と符号化データ位置と撮影位置と光源位置を記述した参照関係テーブルを作成し、参照関係テーブルと符号化データをデータ記録・伝送部（３９１０）に出力する。データ記録・伝送部（３９１２）は、用途に応じて記録・伝送する。

【０１８８】以上のように構成された本実施の形態の画像符号化装置の動作を説明する。撮影位置光源位置制御型画像撮像部（３９０１）では、光源位置が対象物体の周囲を網羅するように、異なる光源位置で画像を撮影する。計算機（３９０４）によって、サーボモータ２（３９０４）を対象物体の正面から水平方向にρ１度、水平面から垂直方向にθ１度の位置に光源（３９０５）を移動させる。さらに、それぞれの光源位置について、カメラ位置が対象物体の周囲を網羅するように、異なる視点から画像を撮影する。

【０１８９】計算機（３９０４）によってサーボモータ１（３９０２）を対象物体の正面から水平方向にρｃ度、水平面から垂直方向にθｃ度の位置にカメラ（３９０３）を移動させ、撮影位置・光源位置・画像データを計算機（３９０４）によって取得する。ここで、得られる画像データは図２と同様に輝度プレーン、色差１プレーン、色差２プレーン、アルファプレーンで構成する。ここで、ρ１，θ１，ρｃ，θｃは、

30

【０１９０】

【数２１】

$$0 \leq \rho_c < 360,\ 0 \leq \theta_c \leq 90,\ (\rho_c, \theta_c \text{は整数})$$

【０１９１】

【数２２】

$$0 \leq \rho_l < 360,\ 0 \leq \theta_l \leq 90,\ (\rho_l, \theta_l \text{は10の倍数})$$

【０１９２】とする。このように取得された撮影位置・光源位置・画像データは、図４０に示すように、各画像データの先頭に水平方向撮影角度ρｃと、垂直方向撮影角度θｃと、水平方向撮影角度ρ１と、垂直方向撮影角度θ１を付加した書式で出力される。

【０１９３】参照予測グループ化部（３９０７）では、参照画像によって類似した画像を予測画像として予測符号化するために、ρ１－θ１－ρｃ－θｃ空間において、等間隔に参照画像を配置し、参照画像の近傍の画像をその参照画像より予測される予測画像にグループ化する。参照予測グループ化部（３９０７）での処理は、図１の参照予測グループ化部（１０５）の処理を４次元的に行う。グループ化された撮影位置・光源位置・画像データはインデックス順に図４１の示す書式で、参照画像は参照画像符号化部（３９０８）へ、予測画像は予測画像符号化部（３９１０）へ出力する。

【０１９４】参照画像符号化部（３９０８）では、参照画像を図１の参照画像符号化部（１０６）と同様に符号化し、符号化データを参照関係テーブル作成部（３９１１）に出力し、参照画像の再構成画像を作成し、再構成参照画像蓄積部（３９０９）に出力する。

【０１９５】再構成参照画像蓄積部（３９０９）では、再構成参照画像を蓄積しておき、予測画像符号化部（３９１０）に出力する。参照予測グループ化部（３９０７）において、図４１のように参照画像と予測画像が整列されて出力されるので、再構成参照画像蓄積部（３９０９）は、画像１枚分の容量を持っていれば十分である。

【０１９６】予測画像符号化部（３９１０）では、図１の予測画像符号化部（１０８）と同様に、予測画像を再構成参照画像を用いて符号化し、符号化データを参照関係テーブル作成部（３９１１）に出力する。参照関係テーブル作成部（３９１１）では、各符号化部より出力された図４２のようなデータを順に走査し、図４３のような参照関係テーブルと図４４のような符号化データを作成し、データ記録・伝送部（３９１２）に出力する。

【０１９７】データ記録・伝送部（３９１２）では、参照関係テーブルと符号化データを磁気ディスクや光ディスクに蓄積しておいたり、ネットワークなどで伝送する。

【０１９８】以上のように、本実施の形態では、異なる

多くの視点から異なる多くの光源位置で撮影された多視点画像を、撮影位置データが近傍の類似した画像である参照画像と１枚以上の予測画像にグループ化し、効率よく符号化することが可能である。さらに、画像を各参照予測グループで参照画像が先頭になるように整列して、符号化手段に出力することで、計算量と計算に必要な蓄積容量を少なくしている。また、その整列順に符号化データ位置を格納したテーブルを付加することで、ランダムアクセスに対応した効率的な復号化が可能である。

【０１９９】図４５は本発明の第１２の実施の形態である画像復号化装置の構成を示すブロック図であり、同図を用いて本実施の形態の構成を説明する。

【０２００】同図において、データ記憶・伝送部（４５０１）では、参照関係テーブルと符号化データを蓄積・伝送し、参照テーブルと符号化データを符号化データ選択部（４５０２）に出力する。符号化データ選択部（４５０２）では、参照関係テーブルから要求された画像のデータの位置を検索し、符号化データから必要なデータを参照画像復号化部（４５０３）や予測画像復号化部（４５０５）に出力する。

【０２０１】要求された光源位置の符号化データがなければ、近傍にある複数の光源位置の符号化データを選択し、それぞれを参照画像復号化部（４５０３）や予測画像復号化部（４５０５）に出力する。参照画像復号化部（４５０３）では、参照画像符号化データを復号化し、再構成参照画像を再構成参照画像蓄積部（４５０４）と中間画像生成部（４５０６）に出力する。再構成参照画像蓄積部（４５０５）は、再構成参照画像を蓄積し、予測画像復号化部（４５０５）に出力する。予測画像復号化部（４５０５）では、予測画像符号化データを復号化し、再構成予測画像を中間画像生成部（４５０６）に出力する。

【０２０２】中間画像生成部（４５０６）では、要求された光源位置の画像がなければ、近傍の光源位置の画像を用いて要求された光源位置の画像の画像を生成し、画像表示部（４５０７）に出力する。画像表示部（４５０７）は入力された画像を蓄積しておき、要求により表示する。

【０２０３】以上のように構成された本実施の形態の画像復号化装置の動作を説明する。データ記憶・伝送部（４５０１）では、参照関係テーブルと符号化データを蓄積・伝送し、参照テーブルと符号化データを符号化データ選択部（４５０２）に出力する。

【０２０４】符号化データ選択部（４５０２）では、要求された光源位置と撮影位置を参照関係テーブルから符号化データ位置を検索し、データ記憶・伝送部（４５０１）から、符号化データ位置の符号化データを取得する。参照／予測フラグと参照符号化データを、参照画像復号化部（４５０３）に出力し、予測符号化データがあれば、予測画像復号化部（４５０５）に出力する。

【０２０５】ただし、参照画像復号化部（４５０３）に出力する参照符号化データが前回と同じ場合は、参照画像はすでに復号化され、蓄積されているので、参照画像復号化部（４５０３）への出力は省略する。また、要求された光源位置のデータがない場合には、光源位置が近傍の４つの符号化データをそれぞれ参照画像復号化部（４５０３）と予測画像復号化部（４５０４）へ出力する。

【０２０６】要求された画像が、（ρ１，θ１，ρｃ，θｃ）であり、参照関係テーブルに、光源位置が水平・垂直方向の角度が１０の倍数の画像しか無い場合、近傍の４つの符号化データ（ρ１１，θ１１，ρｃ，θｃ）、（ρ１１，θ１２，ρｃ，θｃ）、（ρ１２，θ１１，ρｃ，θｃ）、（ρ１２，θ１２，ρｃ，θｃ）は、以下の式によって得られる。

【０２０７】

【数２３】

$$\rho_{l1} = \lceil \rho_l \rceil$$

【０２０８】

【数２４】

$$\rho_{l2} = \lceil \rho_l \rceil + 10$$

【０２０９】

【数２５】

$$\theta_{l1} = \lceil \theta_l \rceil$$

【０２１０】

【数２６】

$$\theta_{l2} = \lceil \theta_l \rceil + 10$$

【０２１１】ここで、「・」は、・を越えない１０の倍数を示す。また、ρ１２が３６０になる場合は、ρ１２＝０とする。

【０２１２】参照画像復号化部（４５０３）では、参照画像が符号化された参照符号化データを復号化し、再構成参照画像を再構成参照画像蓄積部（４５０４）に出力する。

【０２１３】また、参照／予測フラグにおいて、参照画像であれば、中間画像生成部（４５０６）に再構成参照画像を出力する。再構成参照画像蓄積部（４５０４）では、再構成された参照画像を蓄積し、予測画像復号化部（４５０５）に出力する。

【０２１４】予測画像復号化部（４５０５）では、予測画像が符号化された予測符号化データを再構成参照画像を用いて復号化し、中間画像生成部（４５０６）に再構成予測画像を出力する。中間画像生成部（４５０６）では、要求された光源位置と復号化された画像データの光源位置が等しければ、画像表示部（４５０６）へ出力する。等しくなければ、符号化データ選択部（４５０２）で選択された光源位置が近傍の画像より中間画像を生成する。

【０２１５】要求された画像の光源位置を（ρ１，θ１，ρｃ，θｃ）とし、符号化データ選択部（４５０

33

２）で選択された画像の光源位置を（ρｌ１，θｌ１，ρｃ，θｃ）、（ρｌ１，θｌ２，ρｃ，θｃ）、（ρｌ２，θｌ１，ρｃ，θｃ）、（ρｌ２，θｌ２，ρｃ，θｃ）とする。それぞれの近傍画像をＩ１、Ｉ２、Ｉ３、Ｉ４とし、画像Ｉ１の座標（ｘ，ｙ）の画素値を＊

34

＊Ｉ１（ｘ，ｙ）と表わすと、中間画像Ｉの画素値Ｉ（ｘ，ｙ）を以下の式で計算する。

【０２１６】

【数２７】

$$I(x,y) = r_3(r_1 I_1(x,y) + r_2 I_3(x,y) + r_4(r_1 I_2(x,y) + r_2 I_4(x,y)$$

【０２１７】ここで、

【０２１８】

【数２８】

$$r_1 = \frac{\rho_{l2} - \rho_l}{\rho_{l2} - \rho_{l1}}$$

【０２１９】

【数２９】

$$r_2 = \frac{\rho_l - \rho_{l1}}{\rho_{l2} - \rho_{l1}}$$

【０２２０】

【数３０】

$$r_3 = \frac{\theta_{l2} - \theta_l}{\theta_{l2} - \theta_{l1}}$$

【０２２１】

【数３１】

$$r_4 = \frac{\theta_l - \theta_{l1}}{\theta_{l2} - \theta_{l1}}$$

【０２２２】である。生成された中間画像は画像表示部（４５０７）へ出力される。画像表示部（４５０７）では、復号化された画像を蓄積し、表示する。

【０２２３】以上のように、本実施の形態では、多くの異なる視点から多くの異なる光源位置で撮影した画像を、撮影位置が近傍の画像をグループ化し、グループを予測符号化された符号化データを、その参照予測関係と符号化データ位置と撮影位置を記述したテーブルを用いることにより、効率的に復号化することが可能である。

【０２２４】また、参照関係テーブルはグループ順に記述されているので、同一参照予測グループ内が連続的に復号化されるので、再構成参照画像を蓄積しておき、予測符号化データを復号化するだけで、効率的に画像を復号化することが可能である。

【０２２５】さらに、要求された光源位置の画像が符号化されていない場合でも、光源位置が近傍の画像から要求された光源位置の画像を生成するので、要求画像を表示することが可能である。

【０２２６】

【発明の効果】本発明の画像符号化装置と復号化装置では、以下の理由により、従来の画像符号化技術を用いるよりも効率の良い符号化と復号化が可能となる。

【０２２７】１．画像符号化データに参照関係と撮影位置と符号化データ位置を記述したテーブルを付加することにより、参照画像から複数の予測画像に対して予測符号化を用いて符号化する。

【０２２８】２．画像符号化データに参照関係と撮影位置と符号化データ位置を記述したテーブルを参照することにより、要求する画像の符号化データ位置を直ちに検索でき、復号化できる。

【０２２９】３．画像再生の要求を監視しておき、それを参照することにより、高速な要求の変化や計算能力に応じて、自動的に、画質は悪いが計算負荷の小さい変形画像と画質は良いが計算負荷の大きい予測画像を段階的に復号化する。

【０２３０】４．蓄積すべき画像の撮影位置を計算し、復号化処理済みの一枚数の画像を蓄積しておき、復号化処理の回数を減らす。

【０２３１】５．再生すべき画像に変化がないときに、再生すべき画像の撮影位置と撮影位置が復号化する画像撮影位置を計算し、あらかじめ復号化処理しておき、再生すべき画像の変化に高速な応答をする。

【０２３２】６．撮影シーンに動きがない場合に、撮影対象によらず対応点が直線上に拘束できることを利用して、拘束パラメータと１次元的な動き情報で、画像を効率に符号化する。

【０２３３】７．撮影シーンに動きがない場合に、撮影対象によらず対応点が直線上に拘束できることを利用して、拘束パラメータと１次元的な動き情報で、画像を効率に符号化した符号化データを復号化する。

【０２３４】８．画像符号化データに参照関係と撮影位置と光源位置と符号化データ位置を記述したテーブルを付加することにより、多くの視点で多くの照明効果を持つ対象物体画像を、要求にたいして効率的に復号化可能な状態で符号化する。

【０２３５】９．画像符号化データに参照関係と撮影位置と光源位置と符号化データ位置を記述したテーブルが付加された、符号化データを、要求にたいして効率的に復号化する。さらに、要求される光源位置の画像が符号化データに無い場合でも、光源位置が近傍の画像から中間画像を生成する。

【図面の簡単な説明】

【図１】本発明の第１の実施の形態における画像符号化装置のブロック図

【図２】画像データの構成の図

【図３】撮影位置データと画像データの図

【図４】予測符号化の説明図

【図５】参照予測グループごとに整列された撮影位置データと画像データの図

【図６】参照画像符号化部のフローチャート

【図７】予測画像符号化部のフローチャート

【図８】参照予測グループごとに整列された撮影位置データと符号化データの図
【図９】参照関係テーブルの図
【図１０】符号化データの図
【図１１】第２の実施の形態における画像復号化装置のブロック図
【図１２】第３の実施の形態における画像復号化装置のブロック図
【図１３】ランダムアクセステーブルの図
【図１４】ランダムアクセスヘッダの構成の図
【図１５】第４の実施の形態における画像符号化装置のブロック図
【図１６】第５の実施の形態における画像復号化装置のブロック図
【図１７】第６の実施の形態における画像復号化装置のブロック図
【図１８】復号化状況蓄積テーブルの説明図
【図１９】符号化データ選択部（１７０５）の状態テーブルの図
【図２０】符号化データ選択部（１７０５）の出力データテーブルの図
【図２１】第７の実施の形態における画像復号化装置のブロック図
【図２２】符号化データ選択部（２１０４）の状態テーブルの図
【図２３】符号化データ選択部（２１０４）の出力データテーブルの図
【図２４】第８の実施の形態における画像復号化装置のブロック図
【図２５】符号化データ選択部（２４０５）の状態テーブルの図
【図２６】符号化データ選択部（２４０５）の出力データテーブルの図
【図２７】予測テーブルの図
【図２８】予測符号化データ選択部（２４０９）の状態テーブルの図
【図２９】予測符号化データ選択部（２４０９）の出力データテーブルの図
【図３０】第８の実施の形態における画像符号化装置のブロック図
【図３１】エピポーラ拘束の説明図
【図３２】異なる二つの撮影位置の説明図
【図３３】エピポーラパラメータの付加された画像データの図
【図３４】拘束予測符号化部のフローチャート
【図３５】拘束された動き推定方法の説明図
【図３６】拘束予測符号化データの図
【図３７】第９の実施の形態における画像復号化装置のブロック図
【図３８】拘束予測復号化部のフローチャート

【図３９】第１０の実施の形態における画像符号化装置のブロック図
【図４０】撮影位置と光源位置を付加した画像データの図
【図４１】参照予測グループごとに整列された撮影位置データと光源位置データと画像データの図
【図４２】参照予測グループごとに整列された撮影位置データと光源位置データと符号化データの図
【図４３】参照関係テーブル作成部（３９１１）の出力する参照関係テーブルの図
【図４４】参照関係テーブル作成部（３９１１）の出力する符号化データの図
【図４５】第１０の実施の形態における画像復号化装置のブロック図

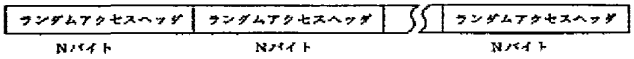
【符号の説明】

１０１　撮影位置制御型画像撮像部
１０２　サーボモータ
１０３　カメラ
１０４　計算機
１０５　参照予測グループ化部
１０６　参照画像符号化部
１０７　再構成参照画像蓄積部
１０８　予測画像符号化部
１０９　参照関係テーブル作成部
１１０　データ記録・伝送部
２０１　輝度プレーン
２０２　色差１プレーン
２０３　色差２プレーン
２０４　アルファプレーン
４０１　参照画像
４０２　予測画像
１１０１　データ記録・伝送部
１１０２　符号化データ選択部
１１０３　参照画像復号化部
１１０４　再構成参照画像蓄積部
１１０５　予測画像復号化部
１１０６　画像表示部
１２０１　データ記録・伝送部
１２０２　ランダムアクセステーブル作成部
１２０３　符号化データ選択部
１２０４　画像要求部
１２０５　参照画像復号化部
１２０６　再構成参照画像蓄積部
１２０７　予測画像復号化部
１２０８　画像表示部
１５０１　撮影位置制御型画像撮像部
１５０２　参照予測グループ化部
１５０３　参照画像符号化部
１５０４　再構成参照画像蓄積部
１５０５　予測画像符号化部

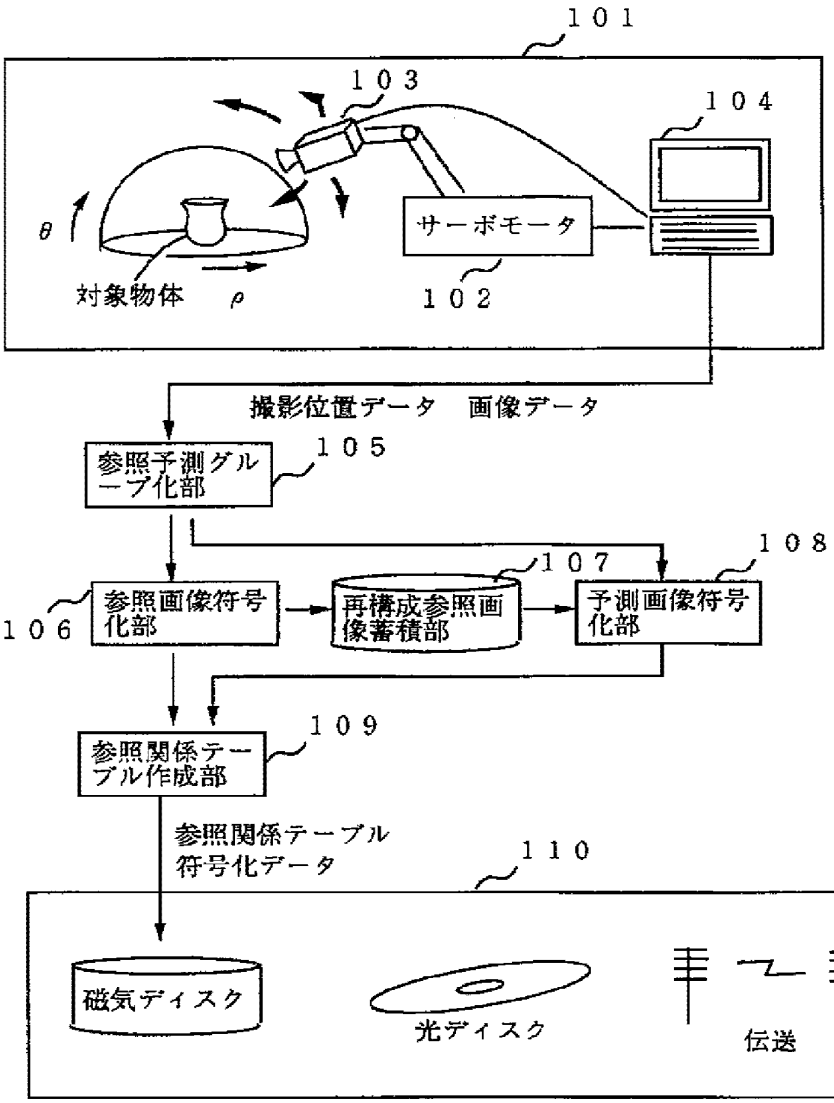
１５０６　参照関係テーブル作成部
１５０７　ランダムアクセステーブル作成部
１５０８　データ記録・伝送部
１６０１　データ記録・伝送部
１６０２　符号化データ選択部
１６０３　画像要求部
１６０４　参照画像復号化部
１６０５　再構成参照画像蓄積部
１６０６　予測画像復号化部
１６０７　画像表示部
１７０１　データ記録・伝送部
１７０２　画像要求監視部
１７０３　画像要求蓄積部
１７０４　復号化状況蓄積部
１７０５　符号化データ選択部
１７０６　参照画像復号化部
１７０７　再構成参照画像蓄積部
１７０８　変形参照画像復号化部
１７０９　変形参照画像蓄積部
１７１０　予測画像復号化部
１７１１　再構成予測画像蓄積部
１７１２　出力画像選択部
１７１３　画像表示部
１８０１　復号化状況テーブル
１８０２　復号化レベルテーブル
２００１　出力データ書式
２００２　出力データテーブル
２１０１　データ記録・伝送部
２１０２　画像要求部
２１０３　画像蓄積検査部
２１０４　符号化データ選択部
２１０５　参照画像復号化部
２１０６　予測画像復号化部
２１０７　参照予測画像復号化部
２１０８　再構成参照画像群蓄積部
２１０９　再構成予測画像群蓄積部
２１１０　出力画像選択部
２１１１　画像表示部
２４０１　データ記録・伝送部
２４０２　画像要求監視部
２４０３　画像要求蓄積部
２４０４　前回画像要求蓄積部
２４０５　符号化データ選択部
２４０６　参照画像復号化部１
２４０７　予測画像復号化部１

２４０８　参照予測画像復号化部１
２４０９　予測符号化データ選択部
２４１０　参照画像復号化部２
２４１１　予測画像復号化部２
２４１２　参照予測画像復号化部２
２４１３　再構成参照画像群蓄積部
２４１４　再構成予測画像群蓄積部
２４１５　画像表示部
３００１　撮影位置制御型画像撮像部
３００２　サーボモータ
３００３　カメラ
３００４　計算機
３００５　参照予測グループ化部
３００６　参照画像符号化部
３００７　再構成参照画像蓄積部
３００８　変形拘束式推定部
３００９　拘束予測画像符号化部
３０１０　参照関係テーブル作成部
３０１１　データ記録・伝送部
３７０１　データ記録・伝送部
３７０２　符号化データ選択部
３７０３　参照画像復号化部
３７０４　再構成参照画像蓄積部
３７０５　拘束予測画像復号化部
３７０６　画像表示部
３９０１　撮影位置光源位置制御型画像撮像部
３９０２　サーボモータ１
３９０３　カメラ
３９０４　サーボモータ２
３９０５　光源
３９０６　計算機
３９０７　参照予測グループ化部
３９０８　参照画像符号化部
３９０９　再構成参照画像蓄積部
３９１０　予測画像符号化部
３９１１　参照関係テーブル作成部
３９１２　データ記録・伝送部
４５０１　データ記録・伝送部
４５０２　符号化データ選択部
４５０３　参照画像復号化部
４５０４　再構成参照画像蓄積部
４５０５　予測画像復号化部
４５０６　中間画像生成部
４５０７　画像表示部

【図１３】



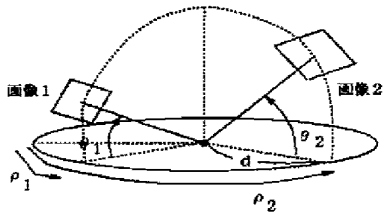
【図１】



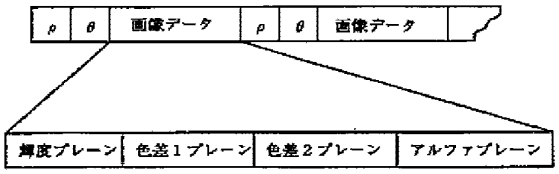
【図２７】

P+(0, 1)
P+(1, 0)
P+(0, −1)
P+(−1, 0)
P+(1, 1)
P+(1, −1)
P+(−1, −1)
P+(−1, 1)

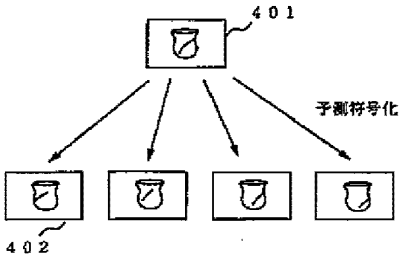
【図３２】



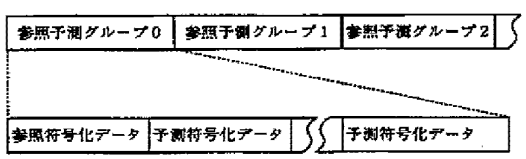
【図３】



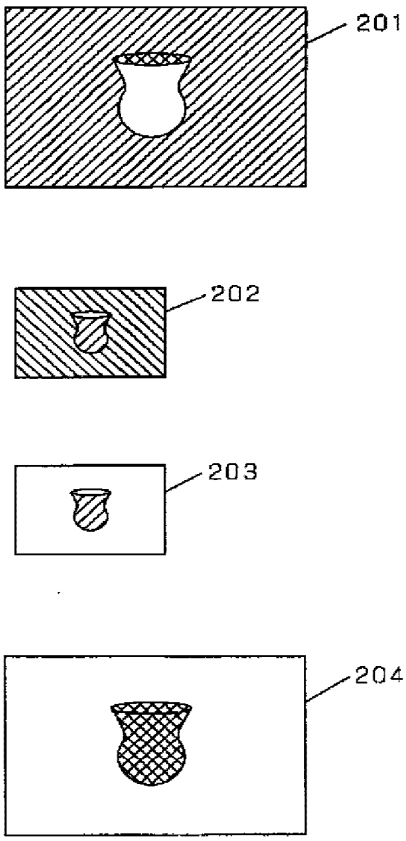
【図４】



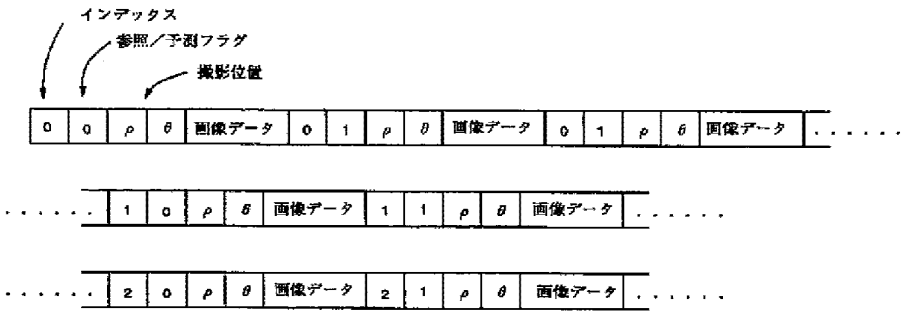
【図１０】



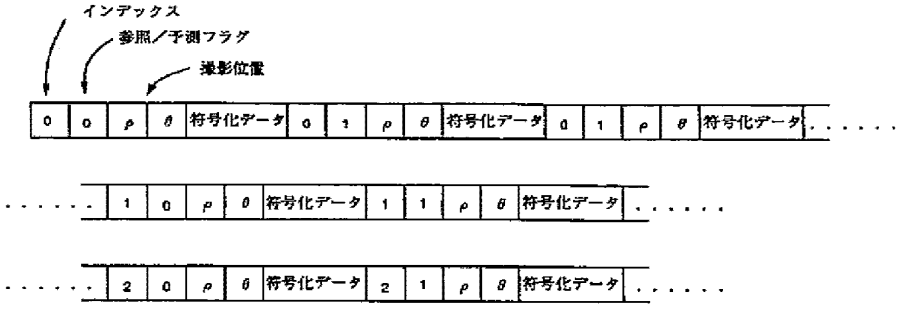
【図２】



【図５】



【図８】



【図１８】

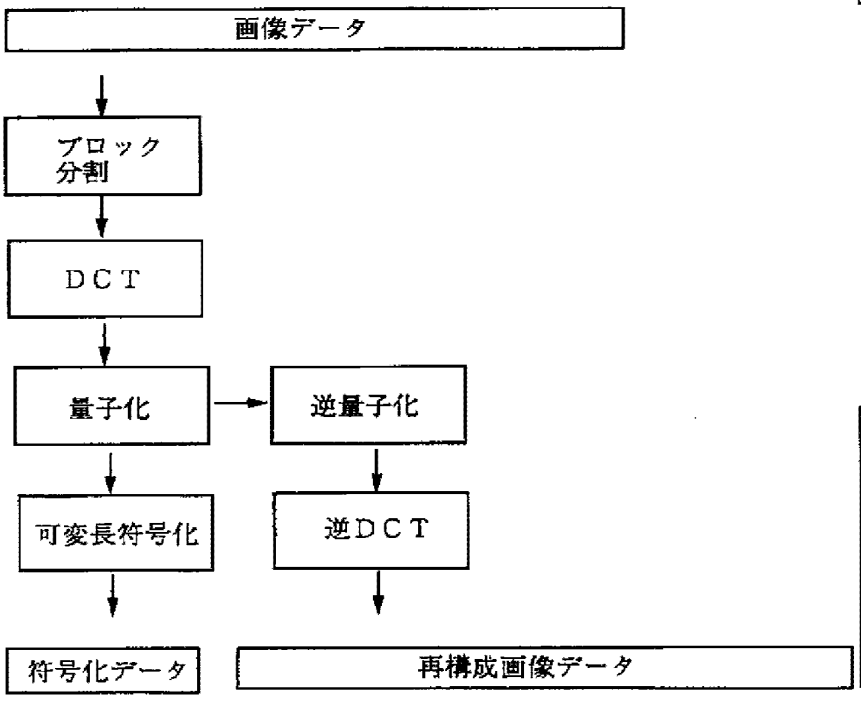
1801

参照画像撮影位置	予測画像撮影位置	復号化レベル

1802

復号化レベル	復号化状況
1	変形参照画像復号化済み
2	予測画像復号化済み

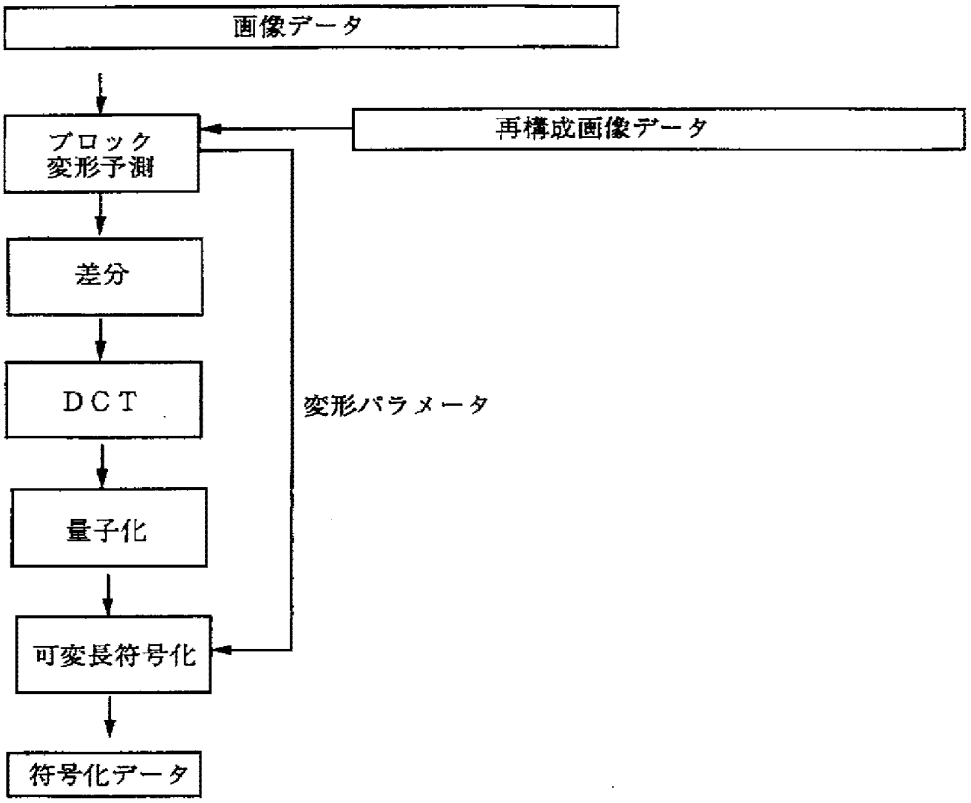
【図６】



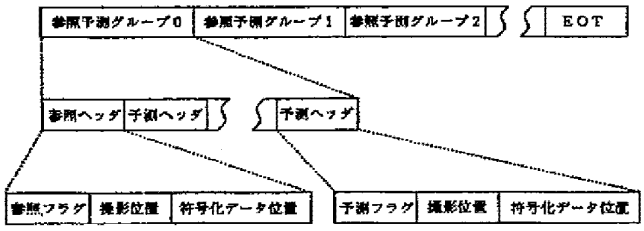
【図１９】

1	位置Ｐは参照画像	Ｐ＝Ｒ		
2		Ｐ≠Ｒ		
3	位置Ｐは予測画像	Ｑ≠Ｒ		
4		Ｑ＝Ｒ	Ｐ≠Ｓ	
5			Ｐ＝Ｓ	Ｌ＝１
6				Ｌ＝２

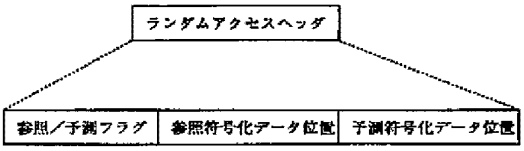
【図７】



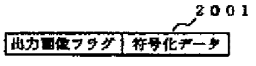
【図９】



【図１４】



【図２０】



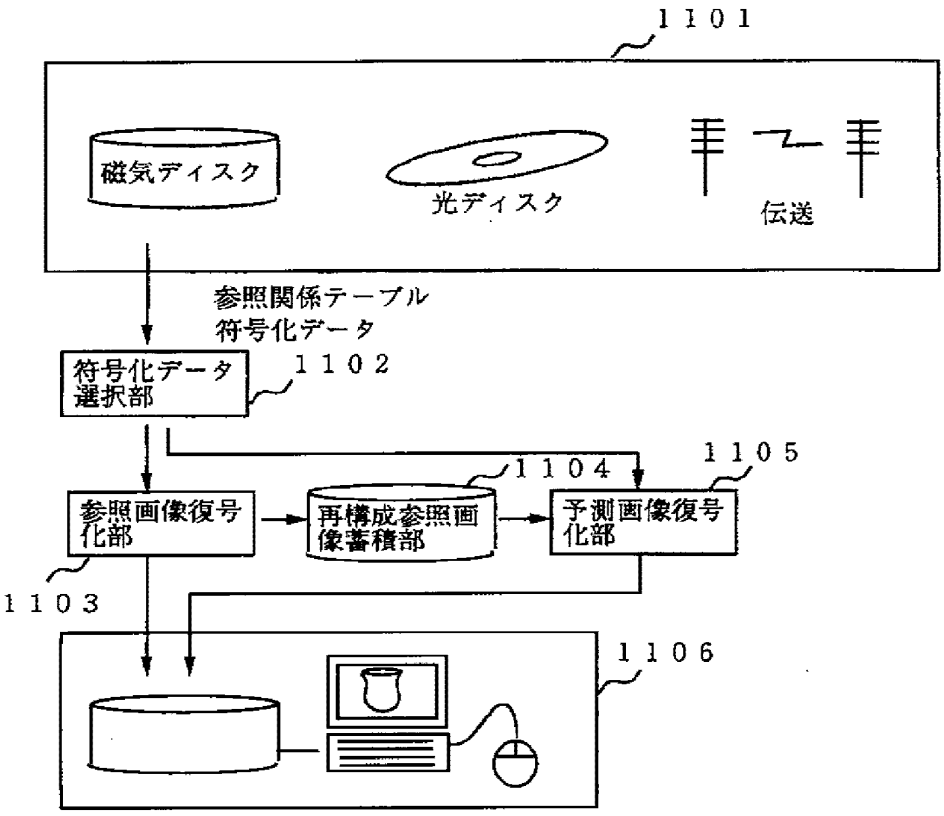
【図２２】

1	位置Ｐは参照画像	Ｐは蓄積されている	
2		Ｐは蓄積されていない	
3	位置Ｐは予測画像	Ｐは蓄積されていない	Ｑは蓄積されている
4			Ｑは蓄積されていない
5		Ｐは蓄積されている	

2002

状態	出力データ	出力先	復号化状況テーブル
1	1	出力画像選択部	R S L
2	1 参照符号化データ（P）	参照画像復号化部	P S L
3	0 参照符号化データ（Q）	参照画像復号化部	Q S L
4	2 予測符号化データ（P）	変形画像復号化部	R P 1
5	3 予測符号化データ（P）	予測画像復号化部	R P 2
6	3	出力画像選択部	R S L

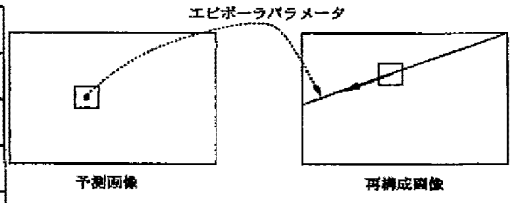
【図１１】



【図２３】

状態	出力先	出力データ
1	出力画像選択部	P
2	参照画像復号化部	P　参照符号化データ（P）
3	予測画像復号化部	P　参照画像データ（Q）　予測符号化データ（P）
4	参照予測画像復号化部	P　参照符号化データ（Q）　予測符号化データ（P）
5	出力画像選択部	P

【図３５】



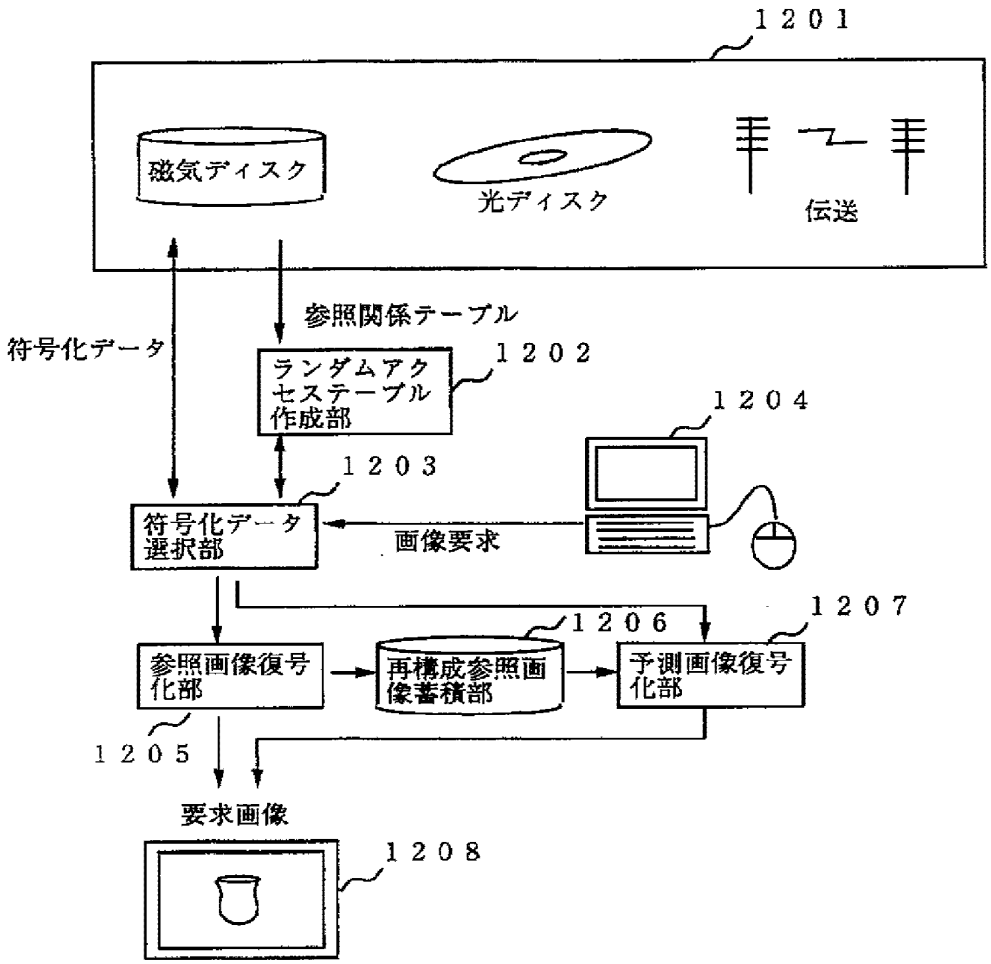
【図２５】

1	P＝R			
2	P≠R	位置Pは参照画像	Pは蓄積されている	
3			Pは蓄積されていない	
4		位置Pは予測画像	Pは蓄積されている	Qは蓄積されている
5				Qは蓄積されていない
6			Pは蓄積されていない	

【図２８】

1	位置Sは参照画像	Sは蓄積されていない	
2	位置Sは予測画像	Sは蓄積されていない	Tは蓄積されている
3			Tは蓄積されていない

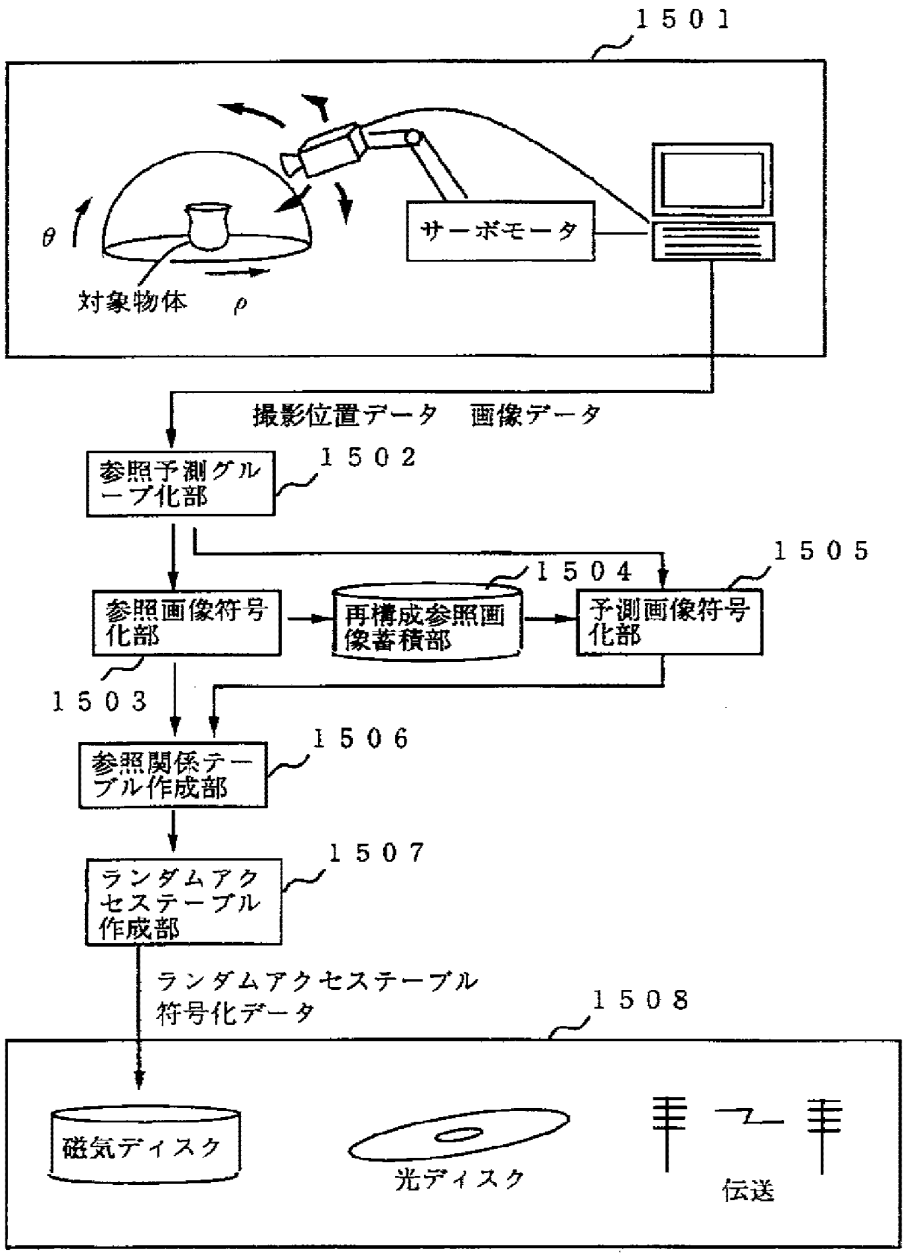
【図１２】



【図２６】

決順	出力先	出力データ
1	予測符号化データ選択部	なし
2	画像表示部	画像データ（P）
3	参照画像復号化部	参照符号化データ（P）
4	予測画像復号化部	参照画像データ（Q）　予測符号化データ（P）
5	参照予測画像復号化部	参照符号化データ（Q）　予測符号化データ（P）
6	画像表示部	画像データ（P）

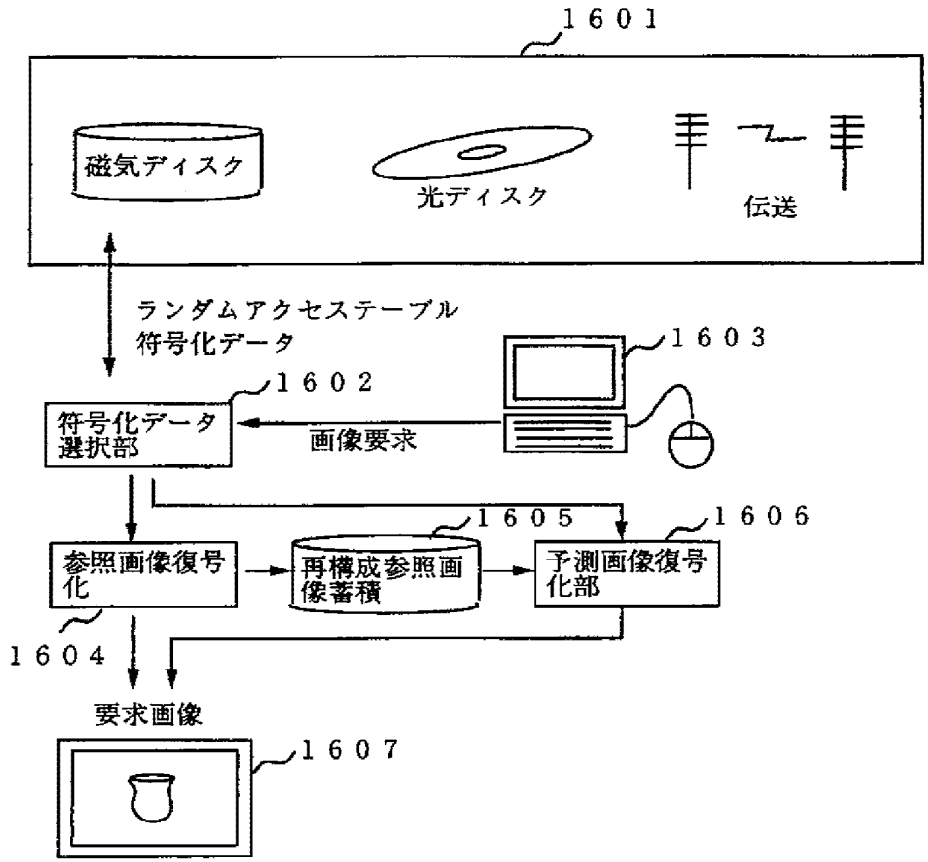
【図１５】



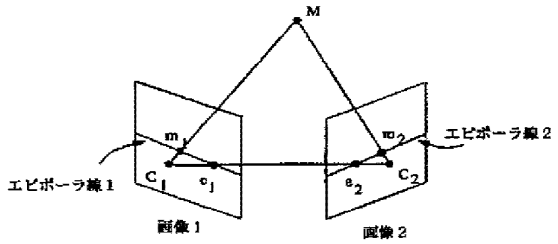
【図２９】

状態	出力先	出力データ
1	参照画像復号化部	参照符号化データ（P）
2	予測画像復号化部	参照画像データ（Q）　予測符号化データ（P）
3	参照予測画像復号化部	参照符号化データ（Q）　予測符号化データ（P）

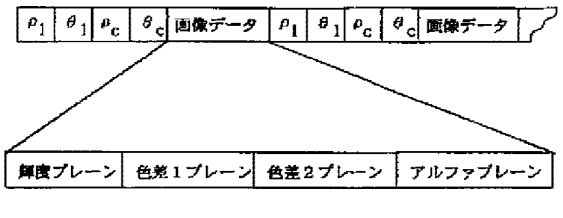
【図１６】



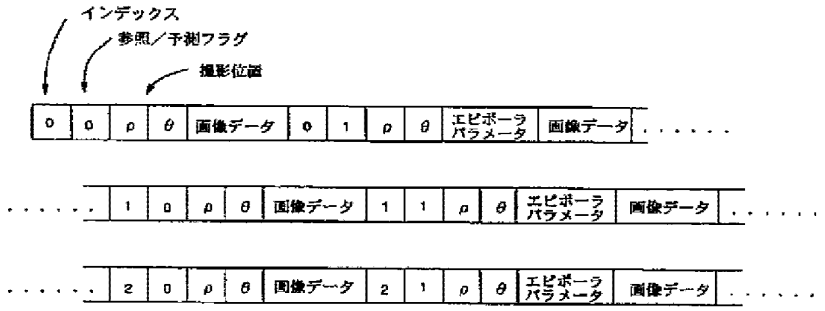
【図３１】



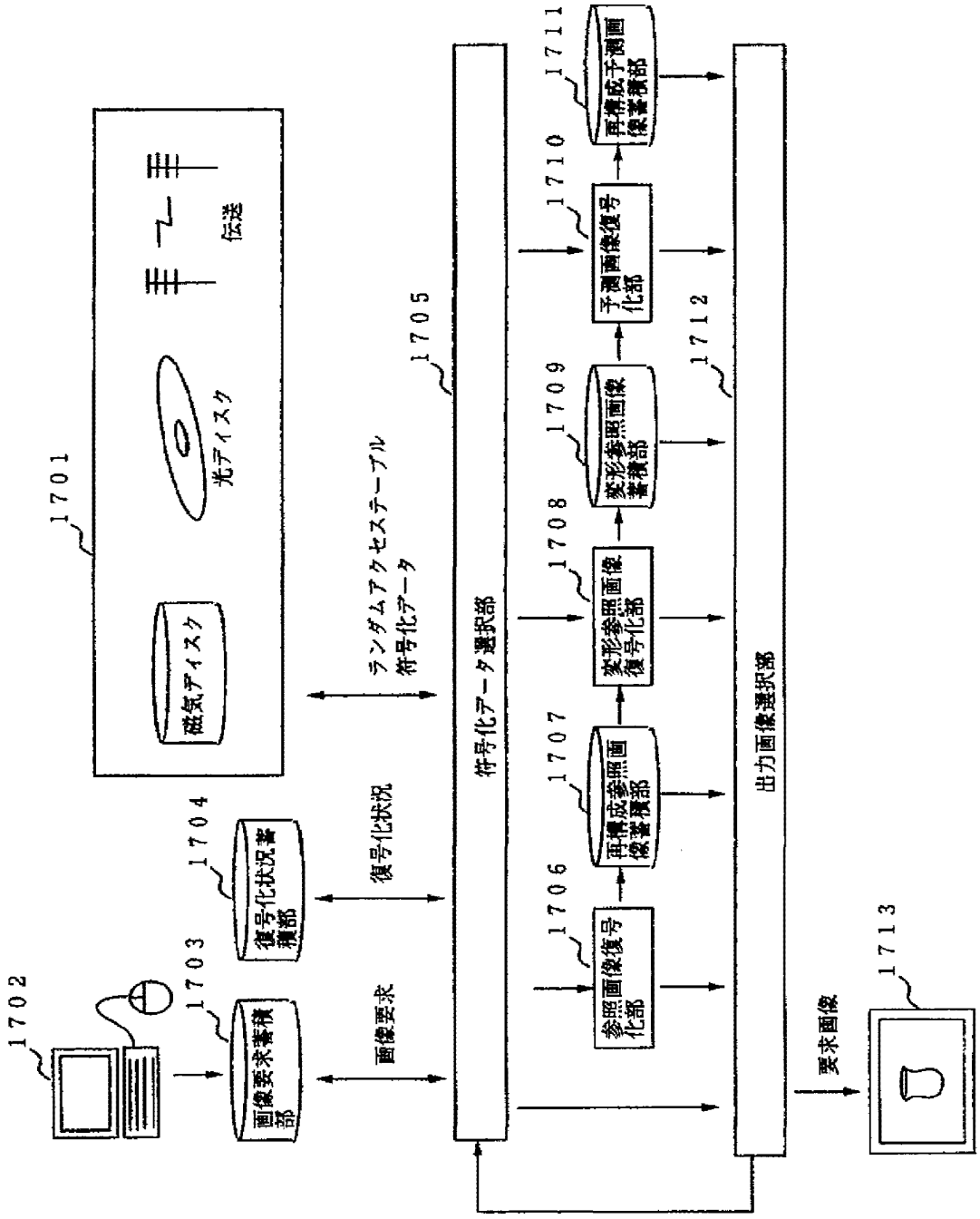
【図４０】



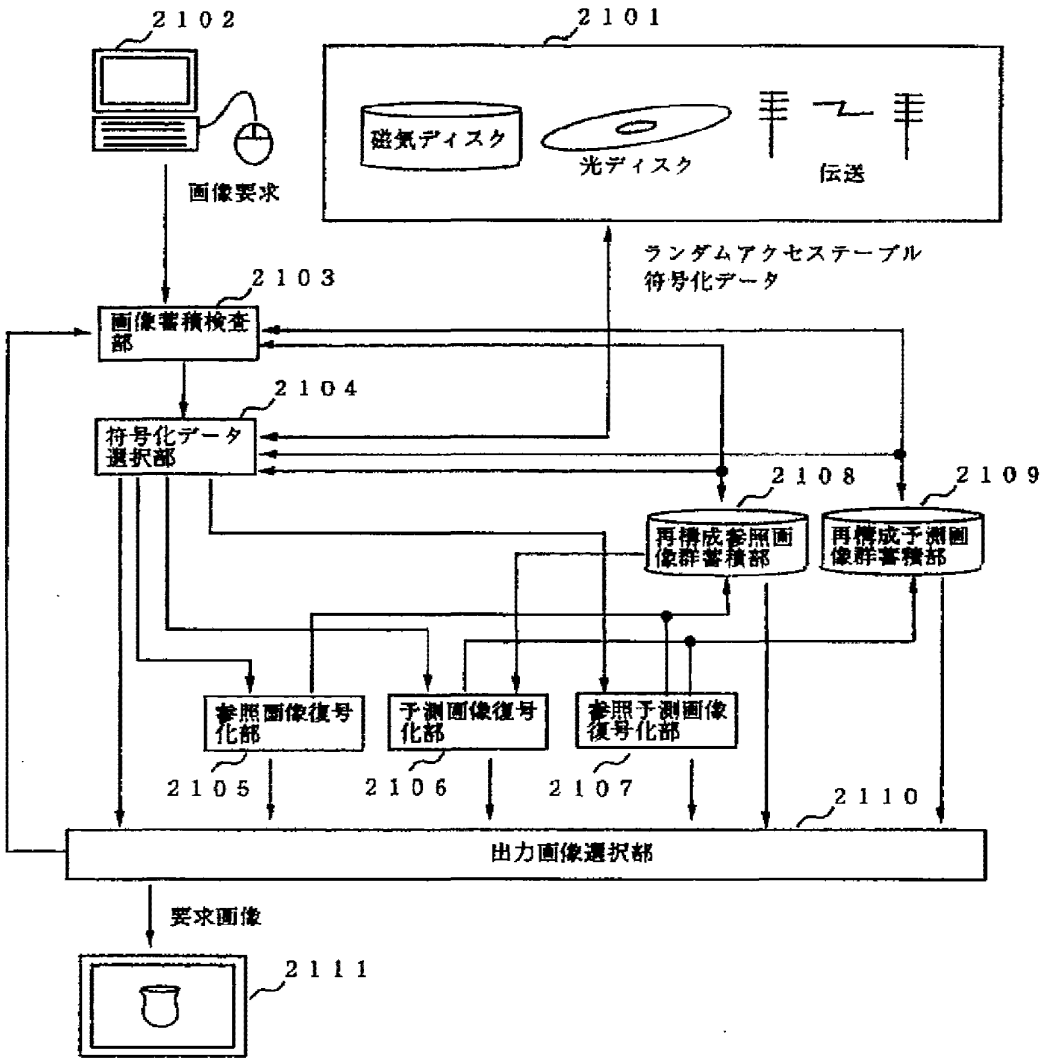
【図３３】



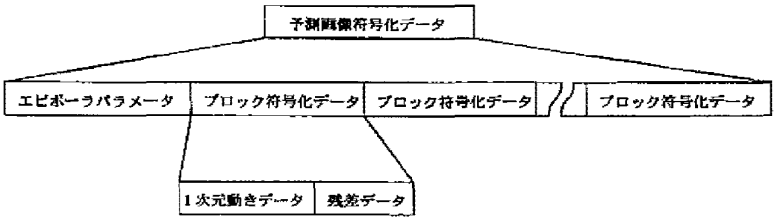
【図１７】


1701
磁気ディスク
光ディスク
伝送
ランダムアクセステーブル
符号化データ
1705
符号化データ選択部
1702
1703
画像要求蓄積部
画像要求
1704
復号化状況蓄積部
復号化状況
1706
参照画像復号化部
1707
再構成参照画像蓄積部
1708
変形参照画像復号化部
1709
変形参照画像蓄積部
1710
予測画像復号化部
1711
画像成予測画像蓄積部
1712
出力画像選択部
1713
要求画像

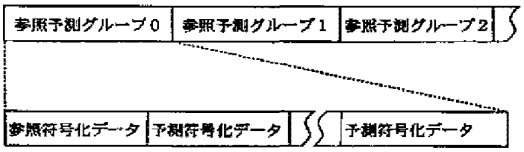
【図２１】


２１０２
２１０１
磁気ディスク
光ディスク
伝送
画像要求
ランダムアクセステーブル
符号化データ
２１０３
画像蓄積検査部
２１０４
符号化データ選択部
２１０８
２１０９
再構成参照画像群蓄積部
再構成予測画像群蓄積部
参照画像復号化部
予測画像復号化部
参照予測画像復号化部
２１０５
２１０６
２１０７
２１１０
出力画像選択部
要求画像
２１１１

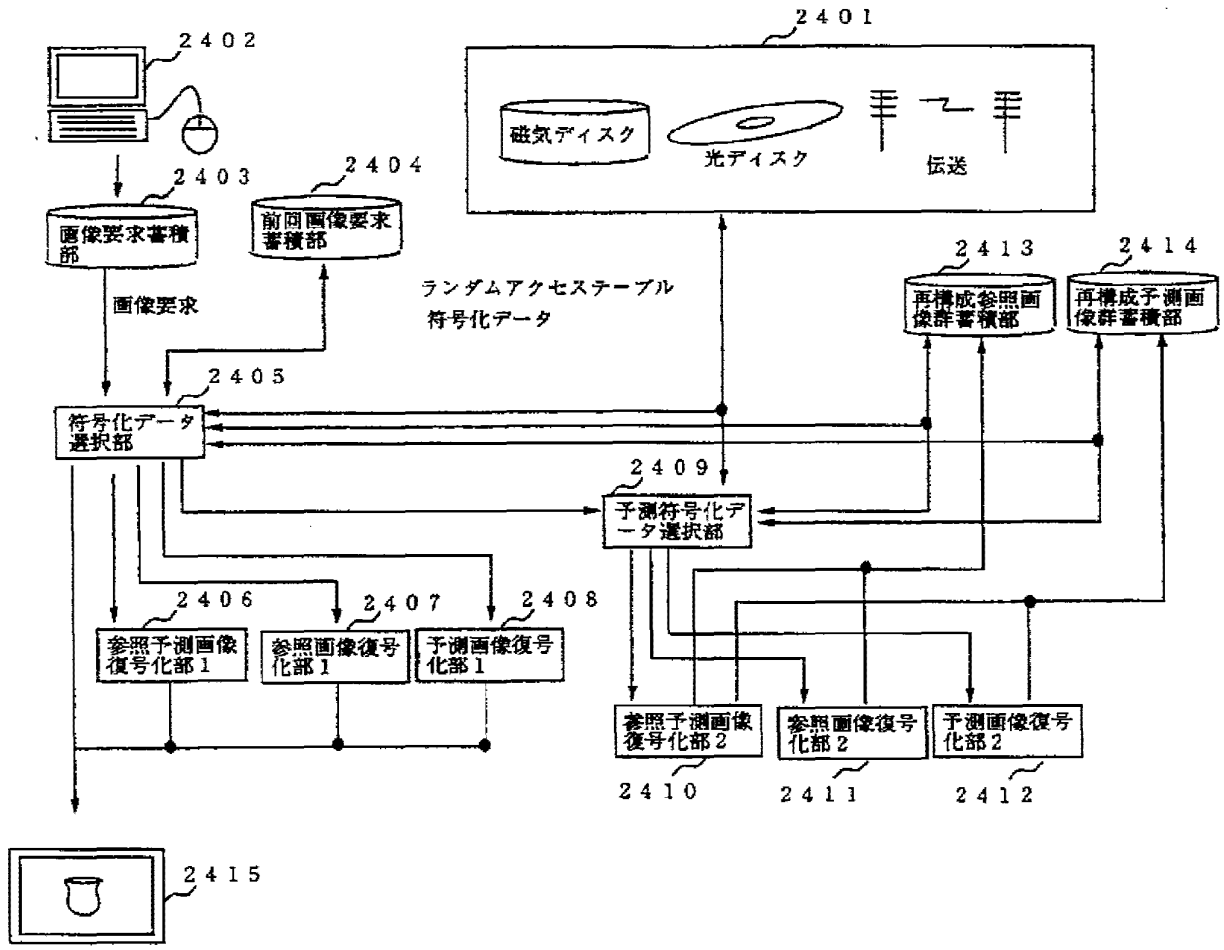
【図３６】


予測画像符号化データ
エピポーラパラメータ
ブロック符号化データ
ブロック符号化データ
ブロック符号化データ
１次元動きデータ
残差データ

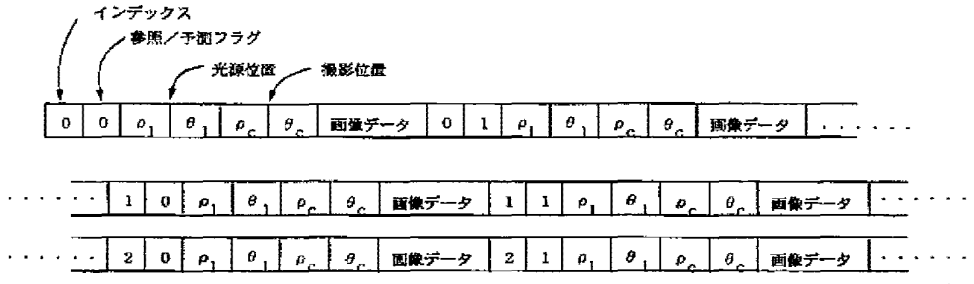
【図４４】


参照予測グループ０
参照予測グループ１
参照予測グループ２
参照符号化データ
予測符号化データ
予測符号化データ

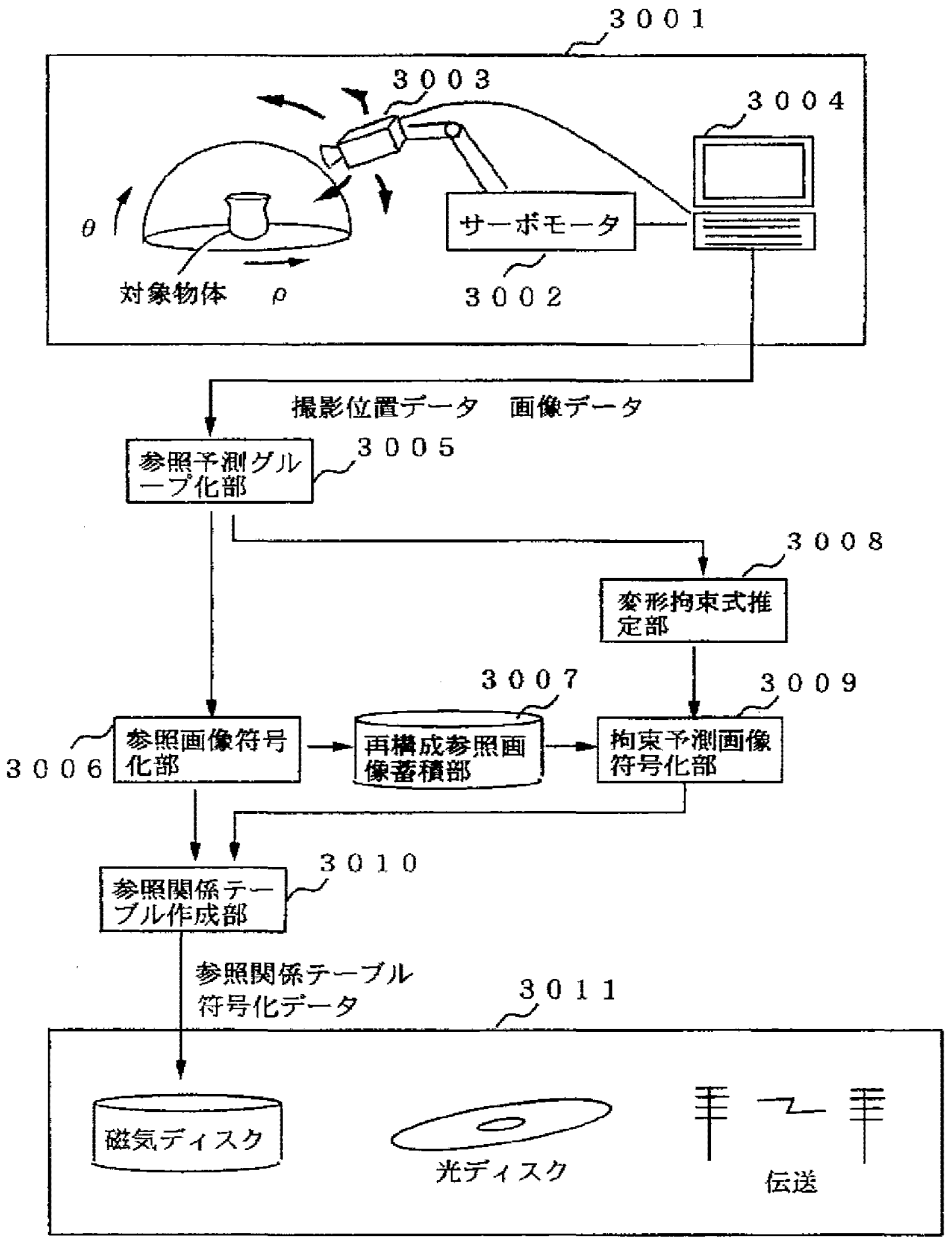
【図２４】


2401
2402
磁気ディスク
光ディスク
伝送
2403
画像要求蓄積部
2404
前回画像要求蓄積部
画像要求
ランダムアクセステーブル
符号化データ
2413
再構成参照画像群蓄積部
2414
再構成予測画像群蓄積部
2405
符号化データ選択部
2409
予測符号化データ選択部
2406
参照予測画像復号化部１
2407
参照画像復号化部１
2408
予測画像復号化部１
参照予測画像復号化部２
参照画像復号化部２
予測画像復号化部２
2410
2411
2412
2415

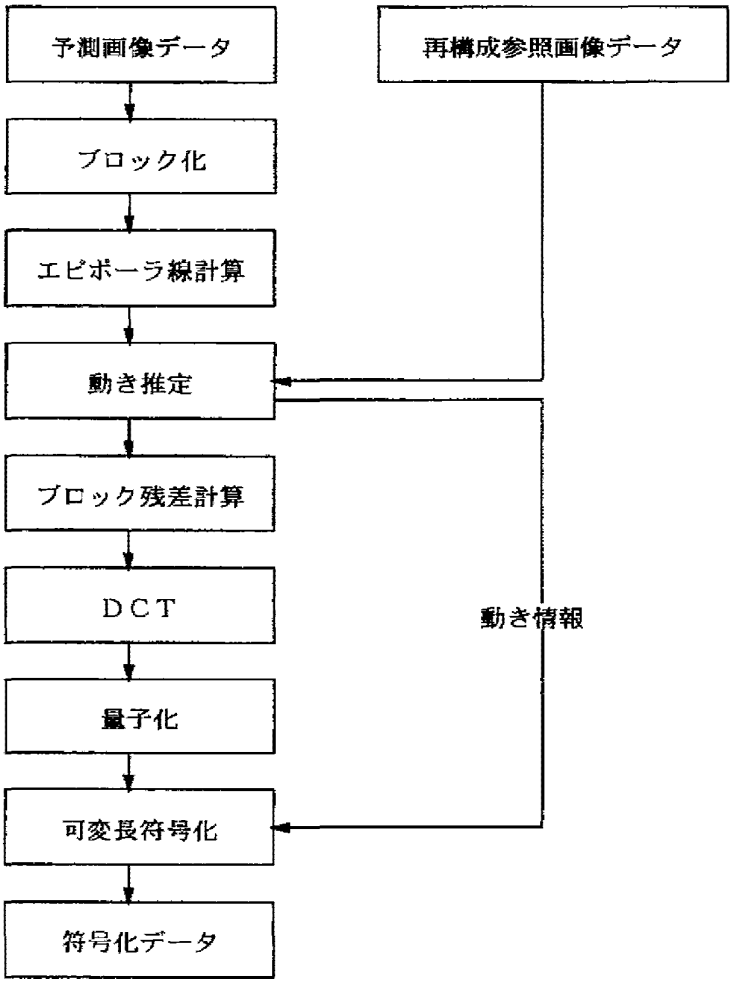
【図４１】


インデックス
参照／予測フラグ
光源位置
撮影位置
画像データ

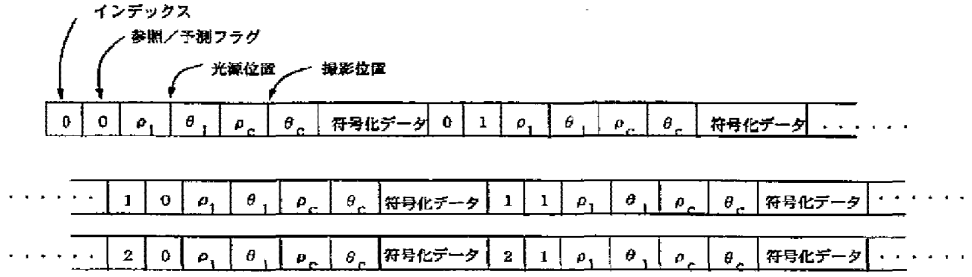
【図３０】


３００１
３００３
３００４
θ
サーボモータ
対象物体　ρ
３００２
撮影位置データ　画像データ
参照予測グループ化部
３００５
３００８
変形拘束式推定部
３００７
３００９
３００６
参照画像符号化部
再構成参照画像蓄積部
拘束予測画像符号化部
３０１０
参照関係テーブル作成部
参照関係テーブル符号化データ
３０１１
磁気ディスク
光ディスク
伝送

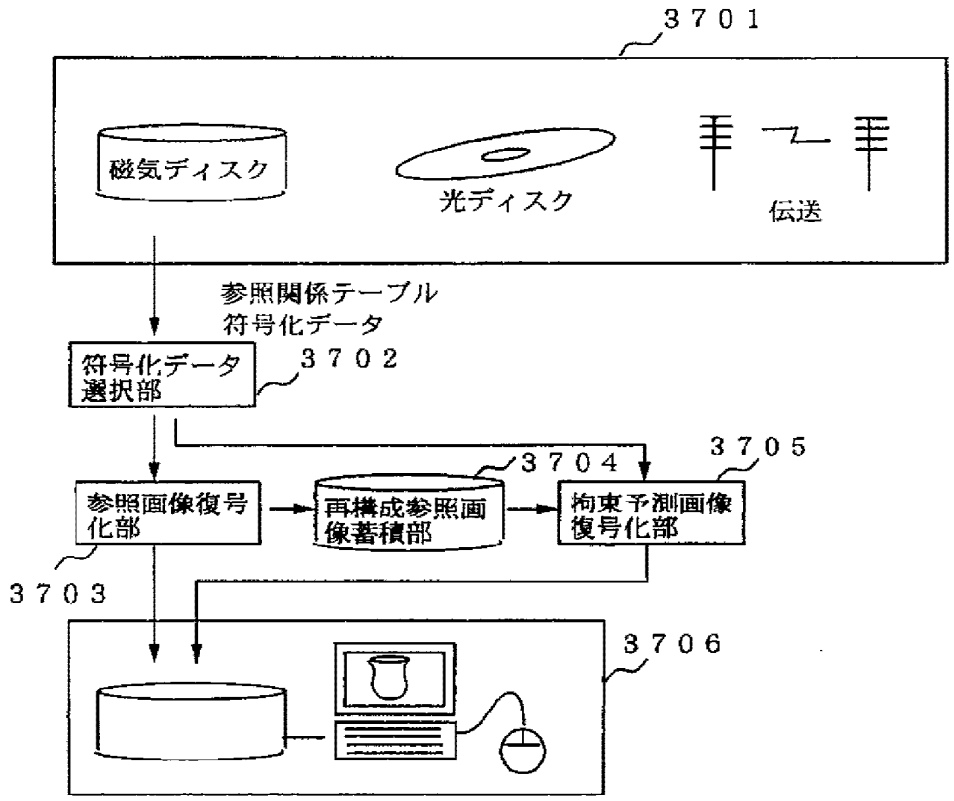
【図３４】


予測画像データ
再構成参照画像データ
ブロック化
エピポーラ線計算
動き推定
ブロック残差計算
ＤＣＴ
動き情報
量子化
可変長符号化
符号化データ

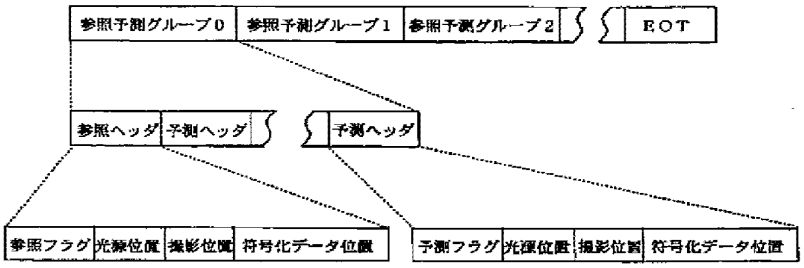
【図４２】


インデックス
参照／予測フラグ
光源位置
撮影位置
符号化データ
符号化データ
符号化データ
符号化データ

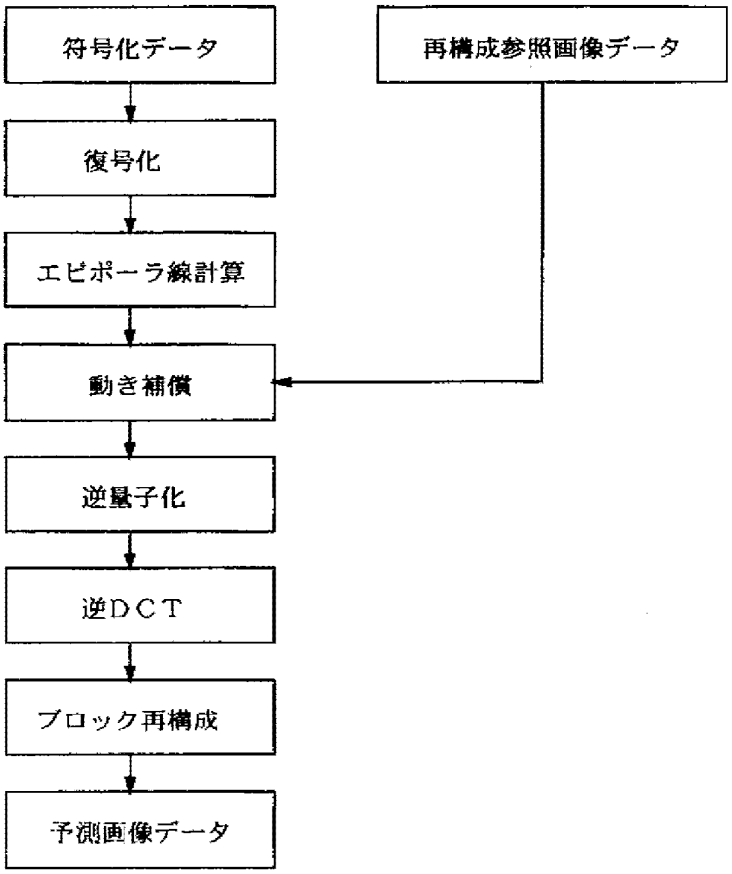
【図３７】


３７０１
磁気ディスク
光ディスク
伝送
参照関係テーブル
符号化データ
符号化データ選択部
３７０２
３７０４
３７０５
参照画像復号化部
再構成参照画像蓄積部
拘束予測画像復号化部
３７０３
３７０６

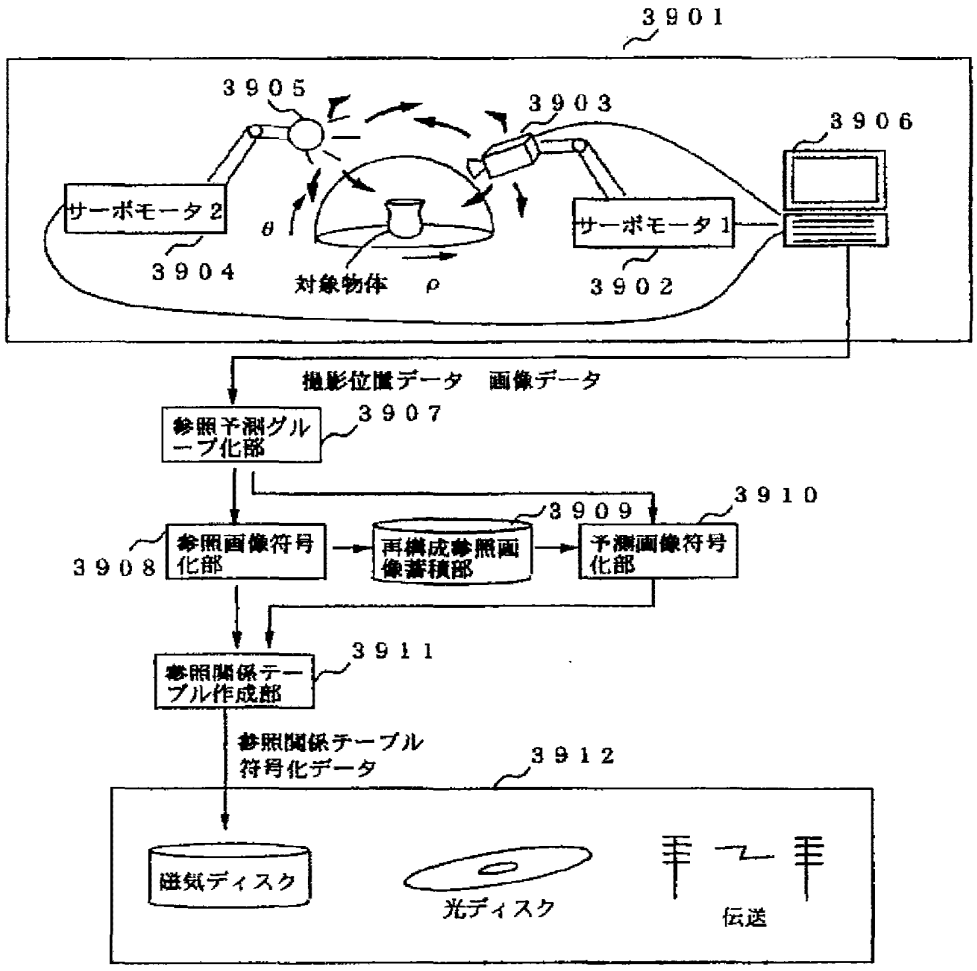
【図４３】


参照予測グループ０
参照予測グループ１
参照予測グループ２
ＥＯＴ
参照ヘッダ
予測ヘッダ
予測ヘッダ
参照フラグ
光源位置
撮影位置
符号化データ位置
予測フラグ
光源位置
撮影位置
符号化データ位置

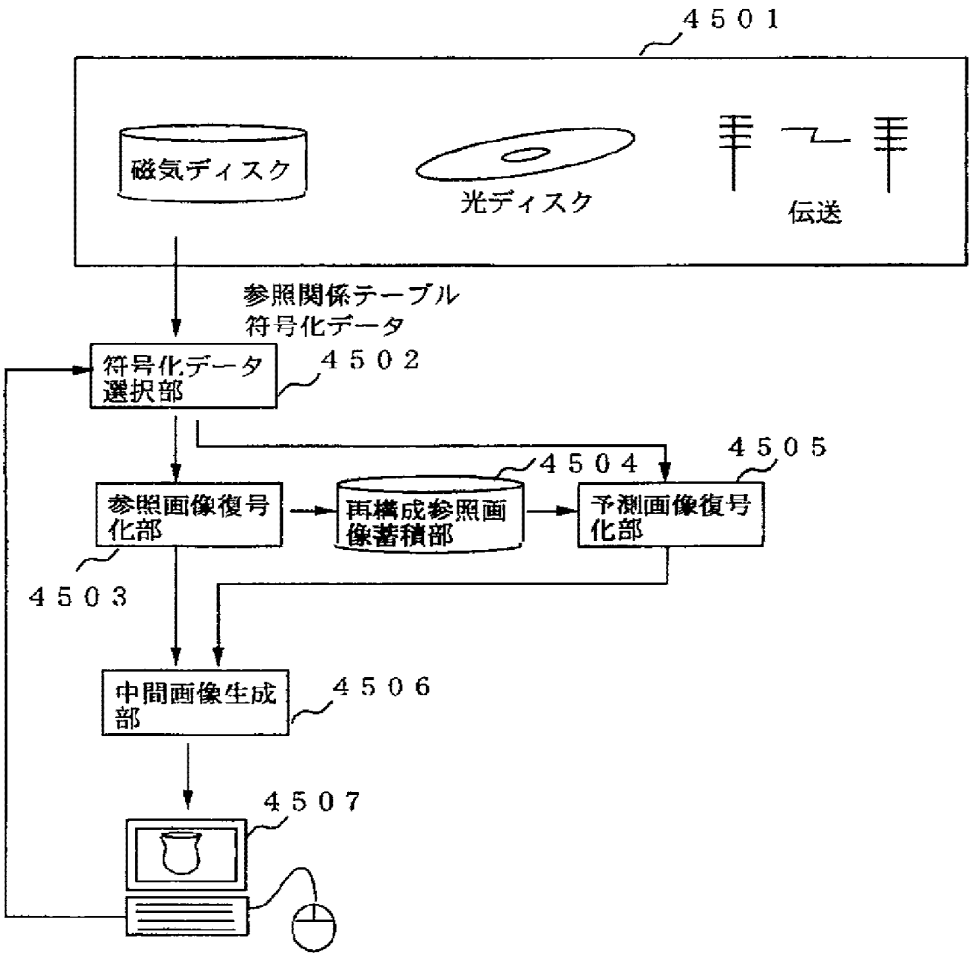
【図３８】


符号化データ
再構成参照画像データ
復号化
エピポーラ線計算
動き補償
逆量子化
逆ＤＣＴ
ブロック再構成
予測画像データ

【図３９】


３９０１
３９０５
３９０３
３９０６
サーボモータ２
θ
サーボモータ１
３９０４
対象物体　ρ
３９０２
撮影位置データ　画像データ
参照予測グループ化部
３９０７
３９１０
３９０９
３９０８
参照画像符号化部
再構成参照画像蓄積部
予測画像符号化部
３９１１
参照関係テーブル作成部
参照関係テーブル符号化データ
３９１２
磁気ディスク
光ディスク
伝送

【図４５】


４５０１
磁気ディスク
光ディスク
伝送
参照関係テーブル
符号化データ
符号化データ選択部
４５０２
４５０４
４５０５
参照画像復号化部
再構成参照画像蓄積部
予測画像復号化部
４５０３
中間画像生成部
４５０６
４５０７

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant:	HANNUKSELA	Examiner:	Unassigned
Serial No.:	11/338,934	Group Art Unit:	2613
Filed:	January 25, 2006	Docket No.:	TPO.004.C2
Confirmation No.:	6610	Customer No.:	76385
Title:	GROUPING OF IMAGE FRAMES IN VIDEO CODING		

SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT (37 C.F.R. § 1.97(b))

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

With regard to the above-identified application, the items of information listed on the enclosed Form 1449 are brought to the attention of the Examiner.

This statement should be considered because it is submitted before the mailing date of a first Office Action on the merits for the above-identified application. Accordingly, no fee is due for consideration of the items listed on the enclosed Form 1449.

The Japanese reference was first cited in a communication from a foreign patent office regarding a counterpart foreign application not more than three months prior to the filing of this information disclosure statement.

In accordance with 37 C.F.R. § 1.98(a)(2), and the 05 August 2003 Official Gazette Notice, only a copy of the foreign document or non-U.S. patent/application listed on the enclosed Form 1449 is provided. The U.S. patent (No. 6,266,158) is included as an English-language version of the cited Japanese reference. Also attached is a copy of the Japanese Office Action with an English translation thereof.

Please note that any notations or markings on the attached documents do not reflect particular relevance, or lack thereof, to the present application, nor were they necessarily made by anyone affiliated with the prosecution of the present application.

No representation is made that a reference is "prior art" within the meaning of 35 U.S.C. §§ 102 and 103 and Applicant reserves the right, pursuant to 37 C.F.R. § 1.131 or otherwise, to establish that the reference(s) are not "prior art." Moreover, Applicant does not represent that a reference has been thoroughly reviewed or that any relevance of any portion of a reference is intended.

Consideration of the items listed is respectfully requested. Pursuant to the provisions of MPEP § 609, it is requested that the Examiner return a copy of the attached Form 1449, marked as being considered and initialed by the Examiner, to the undersigned with the next official communication.

Respectfully submitted,

HOLLINGSWORTH & FUNK, LLC
8009 34th Avenue South, Suite 125
Minneapolis, MN 55425
952.854.2700

By: /Erin M. Nichols/
Erin M. Nichols
Reg. No. 57,125

FORM 1449* INFORMATION DISCLOSURE STATEMENT IN AN APPLICATION (Use several sheets if necessary)	Docket Number: TPO.004.C2	Application Number: 11/338,934
	Applicant: HANNUKSELA	
	Filing Date: 1/25/2006	Group Art Unit: 2613

U.S. PATENT DOCUMENTS

EXAMINER INITIAL	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE

FOREIGN PATENT DOCUMENTS

	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	TRANSLATION	
						YES	NO
	11-177987	02.07.1999	JP			Abstract	

OTHER DOCUMENTS (Including Author, Title, Date, Pertinent Pages, Etc.)

	2010	English Translation of the Office Action of parallel Japanese application no. 2006-120296 dated 01/06/2010.

EXAMINER	DATE CONSIDERED
EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609; draw line through citation if not in conformance and not considered. Include copy of this form for next communication to the Applicant.	

*Substitute Disclosure Statement Form (PTO-1449) Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 11-177987
(43)Date of publication of application : 02.07.1999

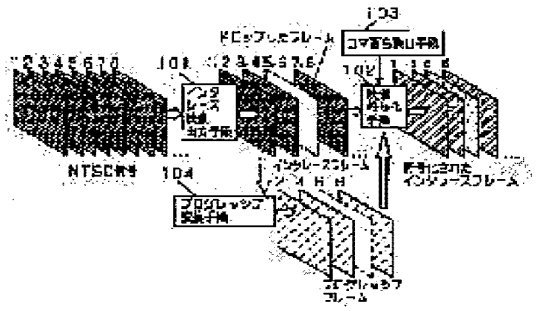
(51)Int.Cl. H04N 7/32

(21)Application number : 09-339675
(22)Date of filing : 10.12.1997

(71)Applicant : MATSUSHITA ELECTRIC IND CO LTD
(72)Inventor : YONEDA AKI
MATSUMOTO TAKAO

(54) VIDEO INFORMATION PROCESSING DEVICE AND METHOD

(57)Abstract:
PROBLEM TO BE SOLVED: To eliminate flickering or fluctuations of images caused when frame dropping occurs in an interlace frame processing and the dropped frame is supplemented by copying a frame or field immediately preceding or following the dropped frame.
SOLUTION: An interlace frame is converted once into a progressive frame 503 which is processed as an interlace frame 501. Thus, half of the time component of the interlace frame is discarded to prevent flickering of a frame-thinned part. The interlace frame is not always converted into a progressive frame but converted only when frame thinning occurs. As a result, the time component loss can be minimized and also computer resources are saved to deal with similar problems.



(19)日本国特許庁（ＪＰ）　(12) **公開特許公報**（Ａ）　(11)特許出願公開番号

特開平11－177987

(43)公開日　平成11年(1999)７月２日

(51)Int.Cl.[6]	識別記号	ＦＩ	
Ｈ０４Ｎ　7/32		Ｈ０４Ｎ　7/137	Ｚ

審査請求　未請求　請求項の数10　ＯＬ　（全 14 頁）

(21)出願番号　特願平9－339675

(22)出願日　平成９年(1997)12月10日

(71)出願人　000005821
松下電器産業株式会社
大阪府門真市大字門真1006番地

(72)発明者　米田　亜旗
大阪府門真市大字門真1006番地　松下電器産業株式会社内

(72)発明者　松本　孝夫
大阪府門真市大字門真1006番地　松下電器産業株式会社内

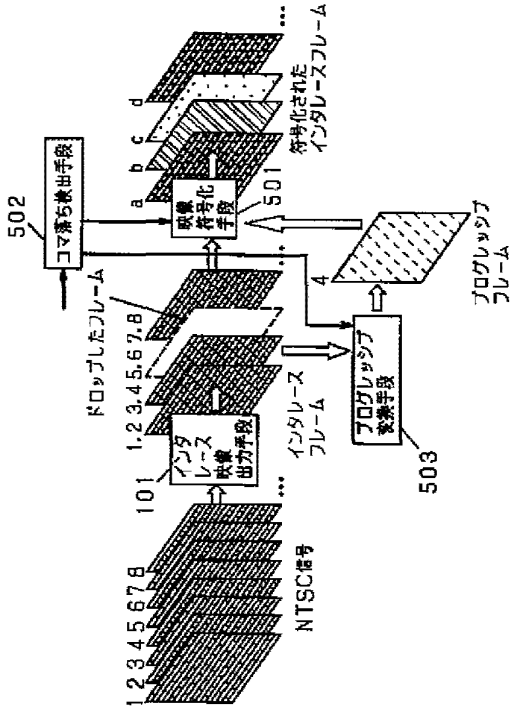
(74)代理人　弁理士　滝本　智之　（外１名）

(54)【発明の名称】　映像情報処理装置及び方法

(57)【要約】

【課題】　インタレースフレームを処理する際に、コマ落ちが発生し、そのコマ落ちを直前あるいは直後のフレームあるいはフィールドをコピーして補う場合に発生する「ちらつき」や「ゆれ」といった問題を解決することを目的とする。

【解決手段】　インタレースフレームを一旦プログレッシブフレームに変換し、そのプログレッシブフレームをインタレースフレームとして処理することで、インタレースフレームが持つ時間成分を半分廃棄し、コマ落ち部分の「ちらつき」を防ぐ。また、常にプログレッシブフレームに変換するのではなく、コマ落ちが発生した時にのみ変換することで、時間成分の損失を最小にとどめ且つ計算機資源を節約して同様の問題に対処する。



1

【特許請求の範囲】

【請求項１】n枚のフィールドから構成されるインタレースフレームを出力するインタレース映像出力手段と、上記n枚のフィールドのうち1枚のフィールドをn倍ライン補間することでプログレッシブフレームに変換するプログレッシブ変換手段と、上記プログレッシブフレームのコマ落ちを検出するコマ落ち検出手段と、上記プログレッシブフレームを、n枚のフィールドから構成されるインタレースフレームとして符号化し、映像符号を出力する符号化手段であり、上記コマ落ち検出手段がコマ落ちを検出すると、コマ落ち部分にはコマ落ち部分の直前あるいは直後の上記プログレッシブフレームと同一の画像であるとする符号を出力し映像符号とする映像符号化手段とを備えたことを特徴とする映像情報処理装置。

【請求項２】n枚のフィールドから構成されるインタレースフレームを出力するインタレース映像出力手段と、上記インタレースフレームのコマ落ちを検出するコマ落ち検出手段と、上記コマ落ち検出手段がコマ落ちを検出すると、コマ落ちの直前あるいは直後の上記インタレースフレームを、上記インタレースフレームを構成するn枚のフィールドのうち1枚のフィールドをn倍ライン補間することでプログレッシブフレームに変換するプログレッシブ変換手段と、上記インタレース映像出力手段が出力したインタレースフレームを符号化し、映像符号を出力する符号化手段であり、上記コマ落ち検出手段がコマ落ちを検出すると、コマ落ちの直前あるいは直後の上記インタレースフレームに代わり対応する上記プログレッシブフレームをn枚のフィールドから構成されるインタレースフレームとして符号化し、且つコマ落ち部分には上記プログレッシブフレームと同一の画像であるとする符号を出力し映像符号とする映像符号化手段とを備えたことを特徴とする映像情報処理装置。

【請求項３】n枚のフィールドから構成されるインタレースフレームを出力するインタレース映像出力手段と、上記インタレースフレームのコマ落ちを検出するコマ落ち検出手段と、上記コマ落ち検出手段がコマ落ちを検出すると、コマ落ちの直前あるいは直後の上記インタレースフレームを構成するn枚のフィールドのうち、時間的にコマ落ち部に近い一枚のフィールドをn倍ライン補間することでプログレッシブフレームに変換した画像を示すダミーフレーム用符号を出力するダミー用符号出力手段と、上記インタレース映像出力手段が出力したインタレースフレームを符号化し、映像符号を出力する符号化手段であり、上記コマ落ち検出手段がコマ落ちを検出すると、コマ落ち部分に出力する符号として、上記ダミーフレーム用符号を出力することで映像符号とする映像符号化手段とを備えたことを特徴とする映像情報処理装置。

【請求項４】n枚のフィールドから構成されるインタレースフレームを出力するインタレース映像出力手段と、

2

上記n枚のフィールドのうち1枚のフィールドをn倍ライン補間することでプログレッシブフレームに変換するプログレッシブ変換手段と、上記プログレッシブフレームのコマ落ちを検出するコマ落ち検出手段と、上記コマ落ち検出手段がコマ落ちを検出すると、コマ落ち部の画像として、コマ落ち部の直前あるいは直後の上記プログレッシブフレームをコピーして補うコマ落ち対処手段と、上記プログレッシブ変換手段及びコマ落ち対処手段とが出力したフレームを、インタレースフレームとして、画像処理を行う画像処理手段とを備えたことを特徴とする映像情報処理装置。

【請求項５】n枚のフィールドから構成されるインタレースフレームを出力するインタレース映像出力手段と、上記インタレースフレームのコマ落ちを検出するコマ落ち検出手段と、上記コマ落ち検出手段がコマ落ちを検出すると、コマ落ちの直前あるいは直後の上記インタレースフレームを構成するn枚のフィールドのうち、時間的にコマ落ち部に近い一枚のフィールドをn倍ライン補間することでプログレッシブフレームに変換し、コマ落ち部の画像を該画像で補うコマ落ち対処手段と、上記インタレース映像出力手段及びコマ落ち対処手段とが出力したフレームを、インタレースフレームとして、画像処理を行う画像処理手段とを備えたことを特徴とする映像情報処理装置。

【請求項６】n枚のフィールドから構成されるインタレースフレームを出力するインタレース映像出力ステップと、上記n枚のフィールドのうち1枚のフィールドをn倍ライン補間することでプログレッシブフレームに変換するプログレッシブ変換ステップと、上記プログレッシブフレームのコマ落ちを検出するコマ落ち検出ステップと、上記プログレッシブフレームを、n枚のフィールドから構成されるインタレースフレームとして符号化し、映像符号を出力する符号化ステップであり、上記コマ落ち検出ステップがコマ落ちを検出すると、コマ落ち部分にはコマ落ち部分の直前あるいは直後の上記プログレッシブフレームと同一の画像であるとする符号を出力し映像符号とする映像符号化ステップとを備えたことを特徴とする映像情報処理方法。

【請求項７】n枚のフィールドから構成されるインタレースフレームを出力するインタレース映像出力ステップと、上記インタレースフレームのコマ落ちを検出するコマ落ち検出ステップと、上記コマ落ち検出ステップがコマ落ちを検出すると、コマ落ちの直前あるいは直後の上記インタレースフレームを、上記インタレースフレームを構成するn枚のフィールドのうち1枚のフィールドをn倍ライン補間することでプログレッシブフレームに変換するプログレッシブ変換ステップと、上記インタレース映像出力ステップが出力するインタレースフレームを符号化し、映像符号を出力する符号化ステップであり、上記コマ落ち検出ステップがコマ落ちを検出すると、コマ

落ちの直前あるいは直後の上記インタレースフレームに代わり対応する上記プログレッシブフレームをn枚のフィールドから構成されるインタレースフレームとして符号化し、且つコマ落ち部分には上記プログレッシブフレームと同一の画像であるとする符号を出力し映像符号とする映像符号化ステップとを備えたことを特徴とする映像情報処理方法。

【請求項８】n枚のフィールドから構成されるインタレースフレームを出力するインタレース映像出力ステップと、上記インタレースフレームのコマ落ちを検出するコマ落ち検出ステップと、上記コマ落ち検出ステップがコマ落ちを検出すると、コマ落ちの直前あるいは直後の上記インタレースフレームを構成するn枚のフィールドのうち、時間的にコマ落ち部に近い一枚のフィールドをn倍ライン補間することでプログレッシブフレームに変換した画像を示すダミーフレーム用符号を出力するダミー用符号出力ステップと、上記インタレース映像を符号化し、映像符号を出力する符号化ステップであり、上記コマ落ち検出ステップがコマ落ちを検出すると、コマ落ち部分に出力する符号として、上記ダミーフレーム用符号を出力することで映像符号とする映像符号化ステップとを備えたことを特徴とする映像情報処理方法。

【請求項９】n枚のフィールドから構成されるインタレースフレームを出力するインタレース映像出力ステップと、上記n枚のフィールドのうち1枚のフィールドをn倍ライン補間することでプログレッシブフレームに変換するプログレッシブ変換ステップと、上記プログレッシブフレームのコマ落ちを検出するコマ落ち検出ステップと、上記コマ落ち検出ステップがコマ落ちを検出すると、コマ落ち部の画像として、コマ落ち部の直前あるいは直後の上記プログレッシブフレームをコピーして補うコマ落ち対処ステップと、上記プログレッシブ変換ステップ及びコマ落ち対処ステップとが出力したフレームを、インタレースフレームとして、画像処理を行う画像処理ステップとを備えたことを特徴とする映像情報処理方法。

【請求項１０】n枚のフィールドから構成されるインタレースフレームを出力するインタレース映像出力ステップと、上記インタレースフレームのコマ落ちを検出するコマ落ち検出ステップと、上記コマ落ち検出ステップがコマ落ちを検出すると、コマ落ちの直前あるいは直後の上記インタレースフレームを構成するn枚のフィールドのうち、時間的にコマ落ち部に近い一枚のフィールドをn倍ライン補間することでプログレッシブフレームに変換し、コマ落ち部の画像を該画像で補うコマ落ち対処ステップと、上記インタレース映像出力ステップ及びコマ落ち対処ステップとが出力したフレームを、インタレースフレームとして、画像処理を行う画像処理ステップとを備えたことを特徴とする映像情報処理方法。

【発明の詳細な説明】

【０００１】

【発明の属する技術分野】本発明は映像情報を符号化する映像情報処理装置に構成に関する。

【０００２】

【従来の技術】近年のデジタル技術の発達により、従来ではアナログの形でしか扱えなかった、ビデオ映像のような映像情報も、デジタルの形で扱えるようになってきた。また、パーソナルコンピュータの高性能化により、デジタル化された映像情報を、パーソナルコンピュータによって処理できるようになってきた。

【０００３】NTSCのようなアナログ映像信号を、一般に販売されているパーソナルコンピュータ用の画像取り込み装置(以下、キャプチャーカード)でデジタルデータに変換したときのデータ構造を、図11で説明する。NTSC信号では、映像は偶数フィールドと、奇数フィールドで構成されている。偶数フィールドと奇数フィールドは、空間的には1ライン分のずれがあり、また時間的には、６０分の１秒分のずれがある。このようなNTSC信号の偶数フィールドと奇数フィールドを、通常のキャプチャーカードでは、奇数フィールドと偶数フィールドがライン毎に組み合わさった、1101のような１枚のフレームとして、３０分の１秒に一枚の間隔で、コンピュータ内に取り込む。このように、本来時間的にも空間的にもずれている二枚のフィールドを、一枚のフレームとして取り扱うとき、そのフレームを、インタレースフレームという。すなわち、現在普及しているキャプチャーカードでは、NTSC信号を、インタレースフレームとしてデジタル化してコンピュータ内に取り込む。

【０００４】上記したようなキャプチャーカードを用い、コンピュータ内に取り込んだインタレースフレームを符号化し、符号化映像情報として出力する映像情報処理装置の従来技術による例を、図を用いて説明する。図12は、従来技術による映像情報処理装置の構成を説明する図である。1201は、インターレース映像出力手段であり、インタレースフレームを出力するもので、一般にはキャプチャーカード等で実現され、NTSC信号を入力し、デジタル化した後にインタレースフレームとして出力するものである。1202は映像符号化手段であり、インタレース映像入力手段(1201)から入力されたインタレースフレームを、符号化し、符号化映像情報とするものである。本例では、インタレースフレームを、そのままインタレースフレームとして、符号化するものとする。例えば、MPEG2(Moving Picture ExpertGroup)の符号化規格によれば、符号化された映像情報が、インタレースフレームであることを示す符号が用意されているので、このような符号化手段を本例は想定している。しかし、それ以外の、例えば、インタレースフレームをフィールドに分割し、フィールド毎に符号化するような符号化手段を適用したとしても、従来技術を説明できる。このことは後に説明する。図13は、従来技術による映像情報処理

装置の動作を説明する図である。

【０００５】上記のように構成された従来技術による映像情報処理装置を、図12と図13を用いて説明する。まず、NTSC信号がインタレース映像出力手段(1201)に入力され、インタレース映像出力手段(1201)は、これをデジタル化し、インタレースフレームとして出力する。映像符号化手段(1202)は、上記インタレースフレームを符号化し、符号化されたインタレースフレームとして、出力する。これを繰り返す。

【０００６】上記のように出力された符号化されたインタレースフレームが復号化される様子を説明するのが、図14である。符号化されたインタレースフレームは、映像復号化手段によって、インタレースフレームに復元される。このインタレーストフレームを、NTSC出力手段が、偶数フィールドと奇数フィールドとに分割し、アナログ信号にして出力する。

【０００７】

【発明が解決しようとする課題】従来技術による映像情報処理装置は、前述の例で説明したように、インタレースフレームを、そのまま符号化するものであった。このような方式は、インタレース映像出力手段が、確実に、すべてのフレームを出力することが条件となる。もし、インタレース映像出力手段が、コマ落としなどをしてしまい、それを符号化しなければならない場合には、従来技術では本質的な問題を有している。

【０００８】図15は、上記の問題を説明するための図である。図15では、６枚のフィールドで構成されるNTSC信号がインタレース映像出力手段に入力されたものとする。６枚のフィールドの時間的な並びを示すため、1～6の番号を示した。1がもっとも過去の映像であり、6がもっとも新しい映像である。上記のごとく６枚のフィールド(1～6)が、インタレース映像出力手段に入力されたが、符号化の負荷が高く、デジタル化したインタレース映像のバッファが足らなくなった、等の理由により、コマ落ちし、６枚のフィールドのうち、中心の二枚(3,4)がインタレースフレームにならなかったものとする。すなわち、本来なら３枚のインタレースフレームが出力されるところが、２枚しか出力されなかった。そこで、映像符号化手段は、コマ落ちしたフレームに対して、「前回のフレームと同一」という符号を出力し(図15のb)、符号化映像情報とした。

【０００９】上記のごとく符号化された映像情報を復号化される様子を説明するのが、図16である。上記で符号化された映像情報は、映像復号化手段によりインタレースフレームに復号化されるが、符号化映像のbは、「前回のフレームと同一」という符号なので、最初の一枚のインタレースフレームが、コピーされ、二枚目のインタレースフレームとして使用される。このように復号化されたインタレースフレームを、NTSC出力手段が、偶数フィールドと奇数フィールドとに分割し、アナログ信号にして出力する。この時、最初のインタレースフレームは、「1,2」という時間的に変化するフィールドを持っており、２枚目のフレームは、最初のフレームを、コピーして使用するので、二枚目のフレームも、「1,2」と時間的に変化するフィールドを持つこととなる。これをNTSC信号として出力すると、最初、「1,2」とフィールドを出力した後に、もう一度、「1」のフィールドを出力する。これは、瞬間的にだが、「過去に戻る」こととなり、動きのある映像を出力すると、「1,2」と動いた映像が「1」で戻るので、映像が「ちらつく」という問題が発生してしまう。

【００１０】図16で示した問題を回避するため、図15におけるbを、「偶数、奇数フィールドとも、前回の偶数フィールドと同じ」という符号とする。すなわち、図16で示した問題は、そもそも時間的に異なる２つのフィールドを持つフレームを、そのままコピーして使用したことに起因するものであるから、単一のフィールドをコピーして使用すれば、問題は回避できるはずである。

【００１１】上記のごとく符号化された映像情報を復号化される様子を説明するのが、図17である。上記で符号化された映像情報は、映像復号化手段によりインタレースフレームに復号化されるが、符号化映像のbは、「偶数、奇数フィールドとも、前回の偶数フィールドと同じ」という符号なので、最初の一枚のインタレースフレームのうち、偶数フィールドのみコピーされ、二枚目のインタレースフレームとして使用される。このように復号化されたインタレースフレームを、NTSC出力手段が、偶数フィールドと奇数フィールドとに分割し、アナログ信号にして出力する。この時、２枚目のフレームは、最初のフレームのうち、偶数フィールドを二回コピーして使用するので、二枚目のフレームは、「2,2」と、時間的に変化しない、同一のフィールドを持つこととなる。これをNTSC信号として出力すると、時間的には、「1,2,2,2,5,6」となり、「過去に戻る」ことはなくなるので、図16のように、「ちらつく」という問題は回避できる。しかし、空間的には、本来、偶数フィールドにある画像を、奇数フィールド、すなわち１ライン上に表示することになるので(図17のX)、画像が上下に「ゆれる」という問題が発生する。これは、静止している映像を出力すると、特に顕著に知覚される問題である。

【００１２】上記した従来の技術による例では、映像符号化手段としては、インタレースフレームを、そのままインタレースフレームとして、符号化するものとしたが、それ以外の、例えば、インタレースフレームをフィールドに分割し、フィールド毎に符号化するような符号化手段を適用したとしても、上記と同様の問題が発生することは、いうまでもない。また、本例では、コマ落ちした部分に対して、「前回のフレームと同じ」あるいは「前回のフィールドと同じ」という符号を用いるものとしたが、符号を用いず、前回のフレーム、あるいはフィ

ールドと同じ映像をそのまま用いてコマ落ちした部分を符号化しても、当然ながら同様の問題が発生する。また、「前回のフレームと同じ」という符号ではなく、「次回のフレームと同じ」という符号を用いたとしても、やはり、同様の問題が発生する。すなわち、コマ落ちした部分を補うのに使用するフィールドが、コマ落ちした部分のフィールドと、偶数、奇数が異なれば、空間的な問題が発生し、偶数、奇数をあわせようとすると、時間的な問題が発生するのである。　本発明は、従来の映像情報処理装置が有していた前記の問題に鑑み、システムの負荷変動によりコマ落ちが発生しても、「ちらつき」や「ゆれ」といった問題を回避しつつ、コマ落ち部分を補うことを可能とすることを目的とするものである。

【００１３】

【課題を解決するための手段】上記目的を達成するために、請求項１、及び請求項４にかかる映像情報処理装置は、すべてのインタレースフレームを、一旦、フレームを構成するフィールドのうちの一枚をプログレッシブ変換することでプログレッシブフレームにし、そのプログレッシブフレームをインタレースフレームとして処理し、コマ落ち部にはコマ落ちの直前あるいは直後のフレームをコピーして補うことを特徴とするものである。

【００１４】また、請求項２にかかる映像情報処理装置は、コマ落ちがあった場合にのみ、コマ落ちの直前あるいは直後のインタレースフレームを、一旦、フレームを構成するフィールドのうちの一枚をプログレッシブ変換することでプログレッシブフレームにし、そのプログレッシブフレームをインタレースフレームとして処理し、コマ落ち部には、コマ落ちの直前あるいは直後のフレームをコピーして補うことを特徴とするものである。

【００１５】また、請求項３及び請求項５にかかる映像情報処理装置は、コマ落ちがあった場合にのみ、コマ落ちの直前あるいは直後のフィールドをプログレッシブ変換することでプログレッシブフレームにし、コマ落ち部には、このプログレッシブフレームをインタレースフレームとして処理することで補うことを特徴とするものである。

【００１６】また、請求項６、及び請求項９にかかる映像情報処理方法は、すべてのインタレースフレームを、一旦、フレームを構成するフィールドのうちの一枚をプログレッシブ変換することでプログレッシブフレームにし、そのプログレッシブフレームをインタレースフレームとして処理し、コマ落ち部にはコマ落ちの直前あるいは直後のフレームをコピーして補うことを特徴とするものである。

【００１７】また、請求項７にかかる映像情報処理方法は、コマ落ちがあった場合にのみ、コマ落ちの直前あるいは直後のインタレースフレームを、一旦、フレームを構成するフィールドのうちの一枚をプログレッシブ変換

することでプログレッシブフレームにし、そのプログレッシブフレームをインタレースフレームとして処理し、コマ落ち部には、コマ落ちの直前あるいは直後のフレームをコピーして補うことを特徴とするものである。

【００１８】また、請求項８及び請求項１０にかかる映像情報処理方法は、コマ落ちがあった場合にのみ、コマ落ちの直前あるいは直後のフィールドをプログレッシブ変換することでプログレッシブフレームにし、コマ落ち部には、このプログレッシブフレームをインタレースフレームとして処理することで補うことを特徴とするものである。

【００１９】

【発明の実施の形態】（実施の形態１）本発明の実施の形態１による映像情報処理装置は、すべてのインタレースフレームのフィールドを一旦プログレッシブフレームに変換し、それをインタレースフレームとして符号化するものである。

【００２０】図1は、本発明の第１の実施の形態における映像情報処理装置の構成図であり、図2は、同装置の動作を説明する図である。図中、101はインターレース映像出力手段であり、インターレースフレームを出力するもので、一般にはキャプチャーカード等で実現され、NTSC信号を入力し、デジタル化した後にインタレースフレームとして出力するものである。102は映像符号化手段であり、プログレッシブ変換手段(104)が変換したプログレッシブフレームを、インタレースフレームとして、符号化するものである。また、コマ落ち検出手段が(103)がプログレッシブ変換手段が出力したプログレッシブフレームのコマ落ちを検出すると、コマ落ち部分に、「コマ落ちの直前のインタレースフレームと同一」という符号を出力するものである。103はコマ落ち検出手段であり、プログレッシブ変換手段(104)が出力するプログレッシブフレームのコマ落ちを検出するものである。104は、プログレッシブ変換手段であり、インタレース映像出力手段(101)が出力したインタレースフレームを構成するフィールドのうち、一枚のフィールドを、プログレッシブフレームに変換するものである。ここで、フィールドをプログレッシブフレームに変換する様子を、図3で示す。図3では、フィールド２の、間隙となっているラインを補間することで、プログレッシブフレームに変換している。フィールド２のピクセルａは、隣あう上下のラインを構成するピクセルの平均をとることで作成している。これをフィールド内補間という。また、フィールド２のピクセルｂは、先立つフィールド１から、対応する位置のピクセルをコピーすることで作成している。これをフィールド間補間という。フィールド間補間をすれば厳密にプログレッシブフレームの解像度が倍になるが、動きのある部分でフィールド間補間をすると、像が２重となり、見ずらい画像となる。フィールド内補間の場合は、解像度は落ちるが、動きのある部分

でも適応できる。このように、フィールドをライン補間してプログレッシブフレームにする技術については、多くの研究がなされている。本例のプログレッシブ変換手段(104)は、いずれかの方法を用いて、インタレースフレームを構成するフィールドのうち、偶数フィールドを、プログレッシブフレームに変換するものとする。上記の例でもわかるように、変換されたプログレッシブフレームは、垂直解像度は倍になっているが、時間的には、偶数フィールドの時点での時間成分のみを持つものとなる。

【００２１】上記のように構成された第１の実施の形態における映像情報処理装置の動作を、図1と図2を用いて説明する。まず、NTSC信号がインタレース映像出力手段(101)に入力され、インタレース映像出力手段(101)は、これをデジタル化し、インタレースフレームとして出力する。ここでは、「1,2,3,4,5,6,7,8」の順序でフィールドが入力されたが、コマ落ちが起こり、本来ならば4枚出力されるはずのインタレースフレームが、「1,2」と、「3,4」と、「7,8」の時間要素を持つ三枚しか、出力されなかった。このインタレースフレームを、プログレッシブ変換手段(104)が、偶数フィールドのみ、プログレッシブフレームにする。偶数フィールドのみをプログレッシブフレームにするので、垂直解像度は倍だが、時間要素は「2」と「4」と「8」である、プログレッシブフレームが３枚出力される。このプログレッシブフレームを、映像符号化手段(102)が、インターレースフレームとして、符号化し、「a」と「b」と「d」を得る。また、コマ落ち検出手段(103)は、プログレッシブフレームに付与されたタイムスタンプを読み取ることで、コマ落ちが発生したことを検出する。コマ落ち検出手段(103)からの指示により、映像符号化手段(102)は、「b」と同一であるという、「c」という符号を出力する。

【００２２】上記のように出力された符号化されたインタレースフレームが復号化される様子を説明するのが、図4である。符号化されたインタレースフレームは、映像復号化手段によって、インタレースフレームに復元される。このインタレーストフレームを、NTSC出力手段が、偶数フィールドと奇数フィールドとに分割し、アナログ信号にして出力する。このとき、「c」は、「b」と同じであるという符号になっており、また、「a」と「b」と「d」は、一度プログレッシブフレームに変換した画像を、インタレースフレームとして符号化したものであるから、垂直解像度はフィールドの倍になっているものの、時間的要素は「2」と「4」と「8」しかない。したがって、再生される映像の時間的要素は、「2,2,4,4,4,4,8,8」となる。すなわち、従来の技術で問題となった、「時間が戻る」ことによる「ちらつき」の問題が解決でき、さらに、プログレッシブフレームに変換することで垂直解像度を補間しているので、本来偶数フィールドで表示すべき画像を奇数フィールドで表示することによる「ゆれ」の問題も解決できる。

【００２３】なお、本例では、偶数フィールドをプログレッシブ変換するものとしたが、これは、奇数フィールドをプログレッシブ変換することもできる。また、本例では、コマ落ち検出手段が、プログレッシブフレームに付与されたタイムスタンプを読み取ることでコマ落ちを検出するものとしたが、プログレッシブ変換手段、あるいはインタレース映像出力手段から、コマ落ちした情報を得ても実現できる。また、本例では、映像符号化手段としては、プログレッシブフレームを、インタレースフレームとして、符号化するものとしたが、それ以外の、例えば、プログレッシブフレームをライン毎に分割してフィールドとし、フィールド毎に符号化するような符号化手段を適用したとしても、本発明は適用できる。また、本例では、コマ落ちした部分に対して、「前回のフレームと同じ」という符号を用いるものとしたが、符号を用いず、前回のフレーム、と同じ映像をそのまま用いてコマ落ちした部分を符号化しても、当然ながら問題は解決できる。また、「前回のフレームと同じ」という符号ではなく、「次回のフレームと同じ」という符号を用いたとしても、やはり、問題は解決できる。

【００２４】さらに、本実施の形態では、出力されたインタレースフレームの処理として、画像の符号化を行うものとしたが、本発明はこのような処理に限定されるものではない。符号化処理ではなく、一般的な画像処理においても、本実施の形態で示したような本発明の実施が行えることはいうまでもない。その場合、コマ落ち部には、「直前、あるいは直後と同一」という符号を用いるのではなく、直前、あるいは直後のフレームをコピーし、コマ落ちを補う、コマ落ち対処手段、等を用意すればよい。

【００２５】（実施の形態２）本発明の実施の形態２による映像情報処理装置は、コマ落ちがあった時に、コマ落ちの直前あるいは直後のインタレースフレームのみ、フィールドを一旦プログレッシブフレームに変換し、それをインタレースフレームとして符号化し、且つ、コマ落ち部分には、プログレッシブフレームに変換した画像と同一であるという符号を出力するものでる。

【００２６】図5は、本発明の第２の実施の形態における映像情報処理装置の構成図であり、図6は、同装置の動作を説明する図である。図中、501は映像符号化手段であり、インタレース映像出力手段(101)が出力したインタレースフレームを符号化するものであり、且つ、コマ落ち検出部(502)がインタレースフレームのコマ落ちを検出すると、コマ落ちの直前のインタレースフレームに関してはこれを符号化せず、代わりにプログレッシブ変換手段(503)が出力したプログレッシブフレームを、インタレースフレームとして、符号化し、さらに、コマ落ち部分に関しては、「直前のインタレースフレームと同一」という符号を出力するものである。502はコマ落

ち検出手段であり、インタレース映像出力手段(101)が出力したインタレースフレームのコマ落ちを検出するものである。503は、プログレッシブ変換手段であり、コマ落ち検出手段(502)がコマ落ちを検出すると、コマ落ち部分の直前のインタレースフレームを構成するフィールドのうち、偶数フィールドを、プログレッシブフレームに変換し、出力するものである。フィールドをプログレッシブフレームに変換する方法については、実施の形態１で説明した方法と同じである。その他の構成要素については実施の形態１と同様であるので、説明は割愛する。

【００２７】上記のように構成された第２の実施の形態における映像情報処理装置の動作を、図5と図6を用いて説明する。まず、NTSC信号がインタレース映像出力手段(101)に入力され、インタレース映像出力手段(101)は、これをデジタル化し、インタレースフレームとして出力する。ここでは、「1,2,3,4,5,6,7,8」の順序でフィールドが入力されたが、コマ落ちが起こり、本来ならば４枚出力されるはずのインタレースフレームが、「1,2」と、「3,4」と、「7,8」の時間要素を持つ三枚しか、出力されなかった。このコマ落ちを、コマ落ち検出手段(502)が、インタレースフレームに付与されたタイムスタンプを読み取ることで検出し、プログレッシブ変換手段(503)に対して、変換の指示を出す。プログレッシブ変換手段(503)は、変換の指示を受けて、コマ落ち部分の直前のインタレースフレームの、偶数フィールドのみ、プログレッシブフレームにする。偶数フィールドのみをプログレッシブフレームにするので、垂直解像度は倍だが、時間要素は「4」である、プログレッシブフレームが一枚出力される。一方、映像符号化手段(501)は、コマ落ち検出手段(502)のコマ落ちの指示を受け、コマ落ち部分の直前のインタレースフレーム(3,4)に関しては符号化せず、代わりにプログレッシブ変換手段(503)が出力したプログレッシブフレームを、インタレースフレームとして、符号化する(図中の「b」)。また、コマ落ち部分に関しては、「直前のインタレースフレームと同一」という符号(図中の「c」)を出力する。さらに、その後のインタレースフレーム(7,8)については、通常どおり、インタレースフレームとして、符号化する(図中の「d」)。

【００２８】上記のように出力された符号化されたインタレースフレームが復号化される様子を説明するのが、図7である。符号化されたインタレースフレームは、映像復号化手段によって、インタレースフレームに復元される。このインタレーストフレームを、NTSC出力手段が、偶数フィールドと奇数フィールドとに分割し、アナログ信号にして出力する。このとき、「c」は、「b」と同じであるという符号になっており、また、「b」は、一度プログレッシブフレームに変換した画像を、インタレースフレームとして符号化したものであるから、垂直解像度はフィールドの倍になっているものの、時間的要素は「4」しかない。したがって、再生される映像の時間的要素は、「1,2,4,4,4,4,7,8」となる。すなわち、従来の技術で問題となった、「ちらつき」の問題が解決でき、さらに、プログレッシブフレームに変換することで垂直解像度を補間しているので、本来偶数フィールドで表示すべき画像を奇数フィールドで表示することによる「ゆれ」の問題も解決できる。さらに、実施の形態１では、すべてのフレームの偶数フィールドをプログレッシブ変換していたため、すべてのフレームの時間的要素が半分となり、コマ落ちが発生しなかったとしても、本来1/60秒に一回動くはずの映像が、1/30秒に一回となり、動きがカクカクするという問題を有していたが、本実施の形態では、コマ落ちが発生した場合にのみ、プログレッシブ変換を行うので、コマ落ちが発生しない限り、1/60秒の滑らかな再生が実現できる。さらに、コマ落ちが生じた時のみプログレッシブ変換を行うことで、毎フレームごとにプログレッシブ変換を行う場合に比べ、計算機資源の節約が図れるという利点もある。

【００２９】なお、本例では、偶数フィールドをプログレッシブ変換するものとしたが、これは、奇数フィールドをプログレッシブ変換することもできる。また、本例では、映像符号化手段としては、プログレッシブフレームを、インタレースフレームとして、符号化するものとしたが、それ以外の、例えば、プログレッシブフレームをライン毎に分割してフィールドとし、フィールド毎に符号化するような符号化手段を適用したとしても、本発明は適用できる。また、本例では、コマ落ちした部分に対して、「前回のフレームと同じ」という符号を用いるものとしたが、符号を用いず、前回のフレーム、と同じ映像をそのまま用いてコマ落ちした部分を符号化しても、当然ながら問題は解決できる。また、コマ落ち部分の直前のインタレースフレームをプログレッシブ変換するのではなく、コマ落ち部分の直後のインタレースフレームをプログレッシブ変換し、コマ落ち部分には、「前回のフレームと同じ」という符号ではなく、「次回のフレームと同じ」という符号を用いても、問題は解決できる。

【００３０】（実施の形態３）本発明の実施の形態３による映像情報処理装置は、コマ落ちがあった時に、コマ落ちの直前あるいは直後のインタレースフレームの、コマ落ち部分に時間的に近いフィールドを、一旦プログレッシブフレームに変換し、それをインタレースフレームとして符号化して、コマ落ち部分の符号として用いるものである。

【００３１】図8は、本発明の第２の実施の形態における映像情報処理装置の構成図であり、図9は、同装置の動作を説明する図である。図中、801は映像符号化手段であり、インタレース映像出力手段(101)が出力したインタレースフレームを符号化するものであり、且つ、コ

マ落ち検出部(502)がインタレースフレームのコマ落ちを検出すると、コマ落ち部分の符号として、ダミー用符号出力手段(802)が出力したダミー用の符号を用いるものである。802は、ダミー用符号出力手段であり、コマ落ち検出部(502)がインタレースフレームのコマ落ちを検出すると、コマ落ち部分の直前のインタレースフレームを構成するフィールドのうち、コマ落ち部に時間的に近いフィールド、すなわち、この場合は偶数フィールドを、プログレッシブフレームに変換し、その画像を符号化して、出力するものである。フィールドをプログレッシブフレームに変換する方法については、実施の形態１で説明した方法と同じである。その他の構成要素については実施の形態２と同様であるので、説明は割愛する。

【００３２】上記のように構成された第３の実施の形態における映像情報処理装置の動作を、図8と図9を用いて説明する。まず、NTSC信号がインタレース映像出力手段(101)に入力され、インタレース映像出力手段(101)は、これをデジタル化し、インタレースフレームとして出力する。ここでは、「1,2,3,4,5,6,7,8」の順序でフィールドが入力されたが、コマ落ちが起こり、本来ならば4枚出力されるはずのインタレースフレームが、「1,2」と、「3,4」と、「7,8」の時間要素を持つ三枚しか、出力されなかった。このコマ落ちを、コマ落ち検出手段(502)が、インタレースフレームに付与されたタイムスタンプを読み取ることで検出し、ダミー用符号出力手段(802)に対して、ダミー用符号出力の指示を出す。ダミー用符号出力手段(802)は、上記の指示を受けて、コマ落ち部分の直前のインタレースフレームの、偶数フィールドを、プログレッシブフレームにする。偶数フィールドのみをプログレッシブフレームにするので、垂直解像度は倍だが、時間要素は「4」である。このプログレッシブフレームを、インタレースフレームとして、符号化して、出力する。一方、映像符号化手段(801)は、コマ落ち検出手段(502)のコマ落ちの指示を受け、コマ落ち部分に、ダミー用符号出力手段(802)が出力したダミー用符号を出力する(図中の「c」)。さらに、その後のインタレースフレーム(7,8)については、通常どおり、インタレースフレームとして、符号化する(図中の「d」)。

【００３３】上記のように出力された符号化されたインタレースフレームが復号化される様子を説明するのが、図10である。符号化されたインタレースフレームは、映像復号化手段によって、インタレースフレームに復元される。このインタレーストフレームを、NTSC出力手段が、偶数フィールドと奇数フィールドとに分割し、アナログ信号にして出力する。このとき、「c」は、一度プログレッシブフレームに変換した画像を、インタレースフレームとして符号化したものであるから、垂直解像度はフィールドの倍になっているものの、時間的要素は「4」しかない。したがって、再生される映像の時間的要素は、「1,2,3,4,4,4,7,8」となる。すなわち、従来

の技術で問題となった、「ちらつき」の問題が解決でき、さらに、プログレッシブフレームに変換することで垂直解像度を補間しているので、本来偶数フィールドで表示すべき画像を奇数フィールドで表示することによる「ゆれ」の問題も解決できる。さらに、実施の形態２では、コマ落ちの直前のフィールドをプログレッシブフレームに変換していたため、コマ落ち部分と合わせて多くの時間成分が失われていたという問題を有していたが、本実施の形態では、コマ落ち部分のみ時間成分が失われるので、滑らかな再生が実現できる。

【００３４】なお、本例では、ダミー用符号出力手段と映像符号化手段としては、プログレッシブフレームを、インタレースフレームとして、符号化するものとしたが、それ以外の、例えば、プログレッシブフレームをライン毎に分割してフィールドとし、フィールド毎に符号化するような符号化手段を適用したとしても、本発明は適用できる。また、本例では、コマ落ちした部分に対して、まず、コマ落ちの直前のフィールドをプログレッシブフレームに変換し、その画像をインタレースフレームとして符号化した符号を用いていたが、本発明を実装する処理系に、「前回のフィールドをプログレッシブ変換したものと同一(あるいは、前回のフィールドのライン補間したものと同一)」という符号が用意されていれば、そのような符号をそのまま用いてもよい。また、コマ落ち部分の直前のフィールドをプログレッシブ変換するのではなく、コマ落ち部分の直後のフィールドをプログレッシブ変換し、それを符号化し、コマ落ち部分に用いても、問題は解決できる。その場合、変換するフィールドは、奇数フィールドとなる。

【００３５】さらに、本実施の形態では、出力されたインタレースフレームの処理として、画像の符号化を行うものとしたが、本発明はこのような処理に限定されるものではない。符号化処理ではなく、一般的な画像処理においても、本実施の形態で示したような本発明の実施が行えることはいうまでもない。その場合、ダミー用符号出力手段の代わりに、直前、あるいは直後のフィールドをプログレッシブ変換し、その画像をコマ落ちを補う、コマ落ち対処手段、等を用意すればよい。

【００３６】さらに、これまで述べた本発明の方法を、プログラム製品として実現し、それをフロッピーディスクやその他の記憶媒体などに格納して配布した場合や、コンピュータネットワーク等の媒体を通じて配布した場合などでも、本発明が有効であることは言うまでもない。

【００３７】

【発明の効果】以上のように、本発明の請求項１にかかる映像情報処理装置は、n枚のフィールドから構成されるインタレースフレームを出力するインタレース映像出力手段と、上記n枚のフィールドのうち1枚のフィールドをn倍ライン補間することでプログレッシブフレームに

15

変換するプログレッシブ変換手段と、上記プログレッシブフレームのコマ落ちを検出するコマ落ち検出手段と、上記プログレッシブフレームを、n枚のフィールドから構成されるインタレースフレームとして符号化し、映像符号を出力する符号化手段であり、上記コマ落ち検出手段がコマ落ちを検出すると、コマ落ち部分にはコマ落ち部分の直前あるいは直後の上記プログレッシブフレームと同一の画像であるとする符号を出力し映像符号とする映像符号化手段とを備えたことを特徴とすることで、インタレースフレームを符号化する際に、コマ落ちが発生し、そのコマ落ちを直前あるいは直後のフレームあるいはフィールドをコピーして補う場合に発生する「ちらつき」や「ゆれ」といった問題を解決することができる。

【００３８】請求項２にかかる映像情報処理装置は、n枚のフィールドから構成されるインタレースフレームを出力するインタレース映像出力手段と、上記インタレースフレームのコマ落ちを検出するコマ落ち検出手段と、上記コマ落ち検出手段がコマ落ちを検出すると、コマ落ちの直前あるいは直後の上記インタレースフレームを、上記インタレースフレームを構成するn枚のフィールドのうち1枚のフィールドをn倍ライン補間することでプログレッシブフレームに変換するプログレッシブ変換手段と、上記インタレースフレームを符号化し、映像符号を出力する符号化手段であり、上記コマ落ち検出手段がコマ落ちを検出すると、上記インタレースフレームに代わり対応する上記プログレッシブフレームをn枚のフィールドから構成されるインタレースフレームとして符号化し、且つコマ落ち部分には上記プログレッシブフレームと同一の画像であるとする符号を出力し映像符号とする映像符号化手段とを備えたことを特徴とすることで、コマ落ちがない部分は1/60fpsの滑らかな動きを保ちつつ、また、プログレッ　シブ変換のための計算機資源を節約した上で、請求項１で述べた効果を実現できる。

【００３９】請求項３にかかる映像情報処理装置は、n枚のフィールドから構成されるインタレースフレームを出力するインタレース映像出力手段と、上記インタレースフレームのコマ落ちを検出するコマ落ち検出手段と、上記コマ落ち検出手段がコマ落ちを検出すると、コマ落ちの直前あるいは直後の上記インタレースフレームを構成するn枚のフィールドのうち、時間的にコマ落ち部に近い一枚のフィールドをn倍ライン補間することでプログレッシブフレームに変換した画像を示すダミーフレーム用符号を出力するダミー用符号出力手段と、上記インタレース映像を符号化し、映像符号を出力する符号化手段であり、上記コマ落ち検出手段がコマ落ちを検出すると、コマ落ち部分に出力する符号として、上記ダミー用符号を出力することで映像符号とする映像符号化手段とを備えたことを特徴とすることで、時間的成分の欠落をコマ落ち部分のみにとどめ、請求項１で述べた効果を実現できる。

16

【００４０】請求項４にかかる映像情報処理装置は、n枚のフィールドから構成されるインタレースフレームを出力するインタレース映像出力手段と、上記n枚のフィールドのうち1枚のフィールドをn倍ライン補間することでプログレッシブフレームに変換するプログレッシブ変換手段と、上記プログレッシブフレームのコマ落ちを検出するコマ落ち検出手段と、上記コマ落ち検出手段がコマ落ちを検出すると、コマ落ち部の画像として、コマ落ち部の直前あるいは直後の上記プログレッシブフレームをコピーして補うコマ落ち対処手段と、上記プログレッシブ変換手段及びコマ落ち対処手段とが出力したフレームを、インタレースフレームとして、画像処理を行う画像処理手段とを備えたことを特徴とすることで、インタレースフレームを処理する際に、コマ落ちが発生し、そのコマ落ちを直前あるいは直後のフレームあるいはフィールドをコピーして補う場合に発生する「ちらつき」や「ゆれ」といった問題を解決することができる。

【００４１】請求項５にかかる映像情報処理装置は、n枚のフィールドから構成されるインタレースフレームを出力するインタレース映像出力手段と、上記インタレースフレームのコマ落ちを検出するコマ落ち検出手段と、上記コマ落ち検出手段がコマ落ちを検出すると、コマ落ちの直前あるいは直後の上記インタレースフレームを構成するn枚のフィールドのうち、時間的にコマ落ち部に近い一枚のフィールドをn倍ライン補間することでプログレッシブフレームに変換し、コマ落ち部の画像を該画像で補うコマ落ち対処手段と、上記インタレース映像出力手段及びコマ落ち対処手段とが出力したフレームを、インタレースフレームとして、画像処理を行う画像処理手段とを備えたことを特徴とすることで、時間的成分の欠落をコマ落ち部分のみにとどめ、請求項４で述べた効果を実現できる。

【００４２】請求項６にかかる映像情報処理方法は、n枚のフィールドから構成されるインタレースフレームを出力するインタレース映像出力ステップと、上記n枚のフィールドのうち1枚のフィールドをn倍ライン補間することでプログレッシブフレームに変換するプログレッシブ変換ステップと、上記プログレッシブフレームのコマ落ちを検出するコマ落ち検出ステップと、上記プログレッシブフレームを、n枚のフィールドから構成されるインタレースフレームとして符号化し、映像符号を出力する符号化ステップであり、上記コマ落ち検出ステップがコマ落ちを検出すると、コマ落ち部分にはコマ落ち部分の直前あるいは直後の上記プログレッシブフレームと同一の画像であるとする符号を出力し映像符号とする映像符号化ステップとを備えたことを特徴とすることで、インタレースフレームを符号化する際に、コマ落ちが発生し、そのコマ落ちを直前あるいは直後のフレームあるいはフィールドをコピーして補う場合に発生する「ちらつき」や「ゆれ」といった問題を解決することができる。

17

【００４３】請求項７にかかる映像情報処理方法は、ｎ枚のフィールドから構成されるインタレースフレームを出力するインタレース映像出力ステップと、上記インタレースフレームのコマ落ちを検出するコマ落ち検出ステップと、上記コマ落ち検出ステップがコマ落ちを検出すると、コマ落ちの直前あるいは直後の上記インタレースフレームを、上記インタレースフレームを構成するn枚のフィールドのうち1枚のフィールドをn倍ライン補間することでプログレッシブフレームに変換するプログレッシブ変換ステップと、上記インタレースフレームを符号化し、映像符号を出力する符号化ステップであり、上記コマ落ち検出ステップがコマ落ちを検出すると、上記インタレースフレームに代わり対応する上記プログレッシブフレームをn枚のフィールドから構成されるインタレースフレームとして符号化し、且つコマ落ち部分には上記プログレッシブフレームと同一の画像であるとする符号を出力し映像符号とする映像符号化ステップとを備えたことを特徴とすることで、コマ落ちがない部分は1/60 fpsの滑らかな動きを保ちつつ、また、プログレッシブ変換のための計算機資源を節約した上で、請求項６で述べた効果を実現できる。

【００４４】請求項８にかかる映像情報処理方法は、ｎ枚のフィールドから構成されるインタレースフレームを出力するインタレース映像出力ステップと、上記インタレースフレームのコマ落ちを検出するコマ落ち検出ステップと、上記コマ落ち検出ステップがコマ落ちを検出すると、コマ落ちの直前あるいは直後の上記インタレースフレームを構成するn枚のフィールドのうち、時間的にコマ落ち部に近い一枚のフィールドをn倍ライン補間することでプログレッシブフレームに変換した画像を示すダミーフレーム用符号を出力するダミー用符号出力ステップと、上記インタレース映像を符号化し、映像符号を出力する符号化ステップであり、上記コマ落ち検出ステップがコマ落ちを検出すると、コマ落ち部分に出力する符号として、上記ダミー用符号を出力することで映像符号とする映像符号化ステップとを備えたことを特徴とすることで、時間的成分の欠落をコマ落ち部分のみにとどめ、請求項６で述べた効果を実現できる。

【００４５】請求項９にかかる映像情報処理方法は、ｎ枚のフィールドから構成されるインタレースフレームを出力するインタレース映像出力ステップと、上記n枚のフィールドのうち1枚のフィールドをn倍ライン補間することでプログレッシブフレームに変換するプログレッシブ変換ステップと、上記プログレッシブフレームのコマ落ちを検出するコマ落ち検出ステップと、上記コマ落ち検出ステップがコマ落ちを検出すると、コマ落ち部の画像として、コマ落ち部の直前あるいは直後の上記プログレッシブフレームをコピーして補うコマ落ち対処ステップと、上記プログレッシブ変換ステップ及びコマ落ち対処ステップとが出力したフレームを、インタレースフレームとして、画像処理を行う画像処理ステップとを備えたことを特徴とすることで、インタレースフレームを処理する際に、コマ落ちが発生し、そのコマ落ちを直前あるいは直後のフレームあるいはフィールドをコピーして補う場合に発生する「ちらつき」や「ゆれ」といった問題を解決することができる。

18

【００４６】請求項１０にかかる映像情報処理方法は、n枚のフィールドから構成されるインタレースフレームを出力するインタレース映像出力ステップと、上記インタレースフレームのコマ落ちを検出するコマ落ち検出ステップと、上記コマ落ち検出ステップがコマ落ちを検出すると、コマ落ちの直前あるいは直後の上記インタレースフレームを構成するn枚のフィールドのうち、時間的にコマ落ち部に近い一枚のフィールドをn倍ライン補間することでプログレッシブフレームに変換し、コマ落ち部の画像を該画像で補うコマ落ち対処ステップと、上記インタレース映像出力ステップ及びコマ落ち対処ステップとが出力したフレームを、インタレースフレームとして、画像処理を行う画像処理ステップとを備えたことを特徴とすることで、時間的成分の欠落をコマ落ち部分のみにとどめ、請求項９で述べた効果を実現できる。

【図面の簡単な説明】

【図１】本発明の第１の実施の形態の映像情報処理装置の構成図

【図２】本発明の第１の実施の形態の映像情報処理装置の動作を説明する図

【図３】プログレッシブ変換を説明する図

【図４】本発明の第１の実施の形態の映像情報処理装置で符号化した映像符号化情報を復号化した時の動作を説明する図

【図５】本発明の第２の実施の形態の映像情報処理装置の構成図

【図６】本発明の第２の実施の形態の映像情報処理装置の動作を説明する図

【図７】本発明の第２の実施の形態の映像情報処理装置で符号化した映像符号化情報を復号化した時の動作を説明する図

【図８】本発明の第３の実施の形態の映像情報処理装置の構成図

【図９】本発明の第３の実施の形態の映像情報処理装置の動作を説明する図

【図１０】本発明の第３の実施の形態の映像情報処理装置で符号化した映像符号化情報を復号化した時の動作を説明する図

【図１１】インタレースフレームを説明する図

【図１２】従来技術による映像情報処理装置の構成図

【図１３】従来技術による映像情報処理装置の動作を説明する図

【図１４】従来技術による映像情報処理装置で符号化した映像符号化情報を復号化した時の動作を説明する図

19　　20

【図１５】従来技術による映像情報処理装置においてコマ落ちが発生した時の動作を説明する図

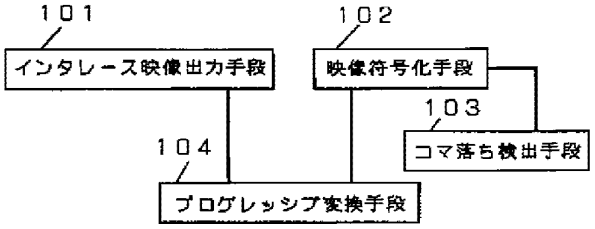
【図１６】従来技術による映像情報処理装置で符号化した映像符号化情報を復号化した際の不具合を説明する第１の図

【図１７】従来技術による映像情報処理装置で符号化した映像符号化情報を復号化した際の不具合を説明する第２の図

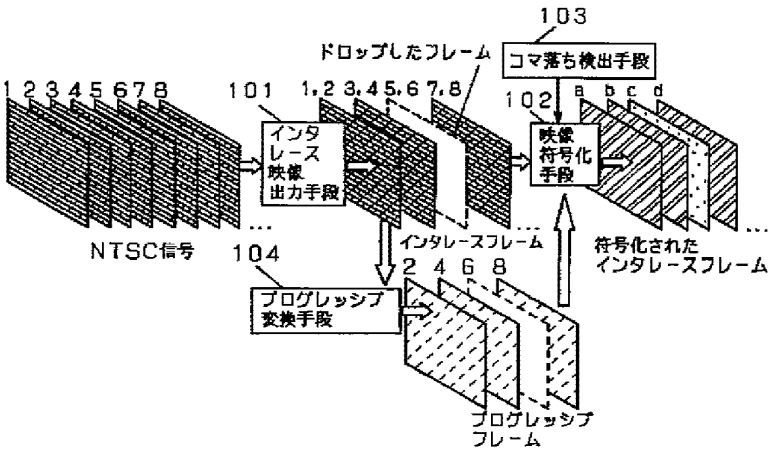
【符号の説明】

１０１　インタレース映像出力手段
１０２　映像符号化手段
１０３　コマ落ち検出手段
１０４　プログレッシブ変換手段

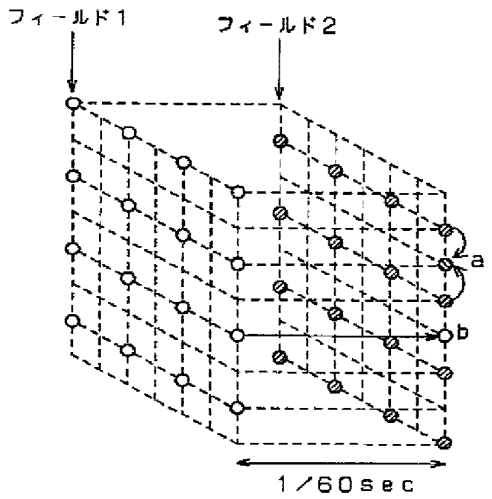
【図１】



【図２】

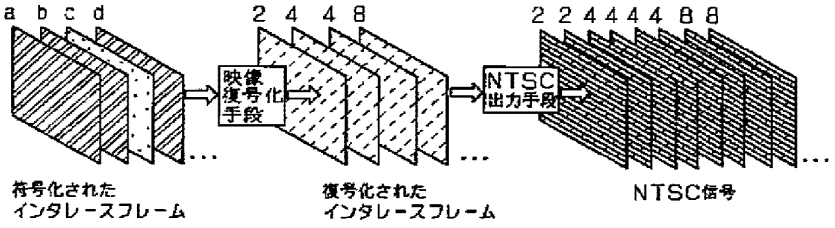


【図３】

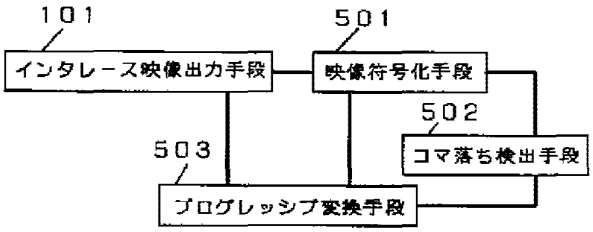


【図１２】

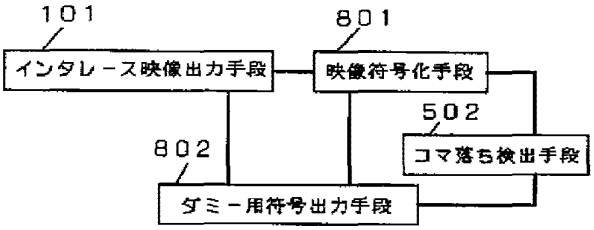
【図４】



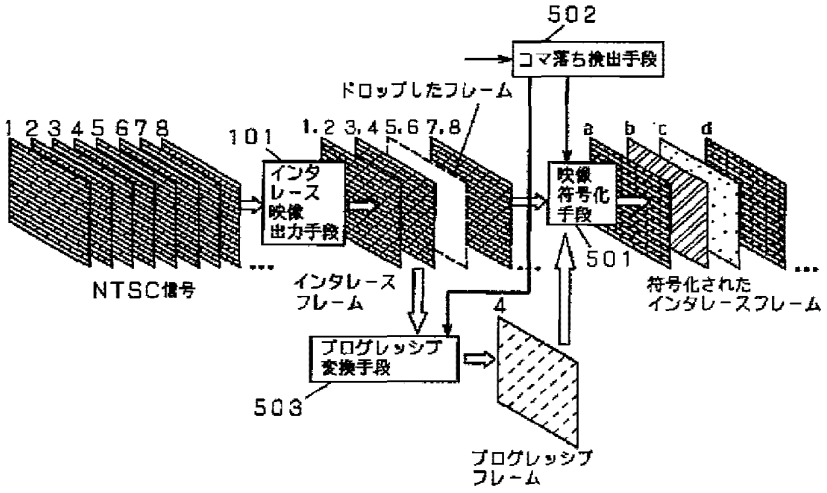
【図５】



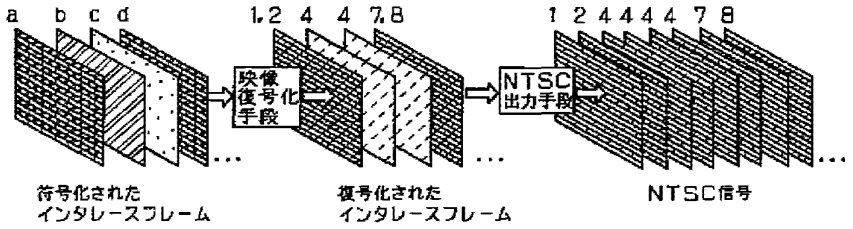
【図８】



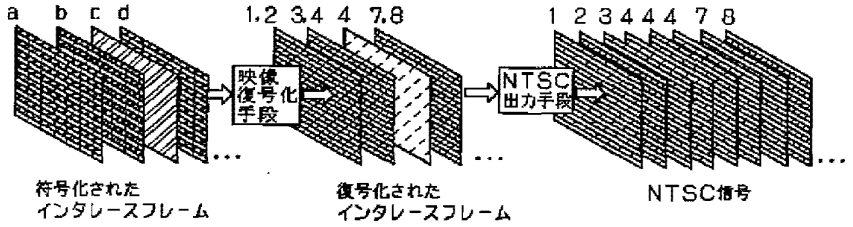
【図６】



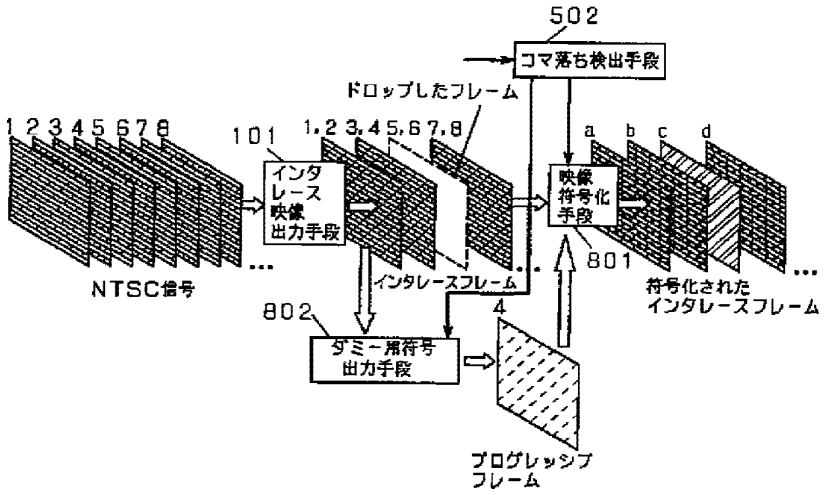
【図７】



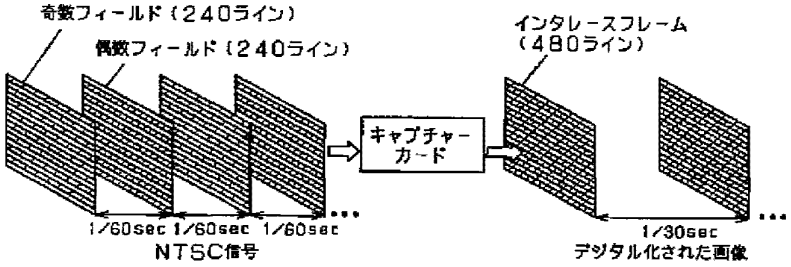
【図１０】



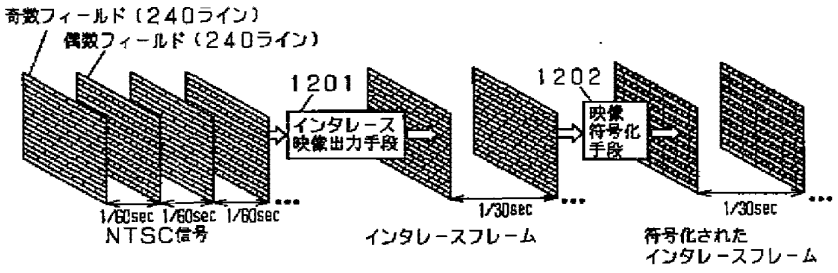
【図９】


502
コマ落ち検出手段
ドロップしたフレーム
1 2 3 4 5 6 7 8
101
インタレース映像出力手段
1 2 3 4 5 6 7 8
a b c d
映像符号化手段
801
NTSC信号
インタレースフレーム
符号化されたインタレースフレーム
802
ダミー用符号出力手段
4
プログレッシブフレーム

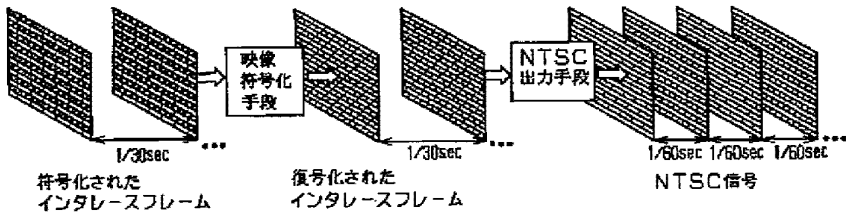
【図１１】


奇数フィールド（240ライン）
偶数フィールド（240ライン）
インタレースフレーム（480ライン）
キャプチャーカード
1/60sec 1/60sec 1/60sec
NTSC信号
1/30sec
デジタル化された画像

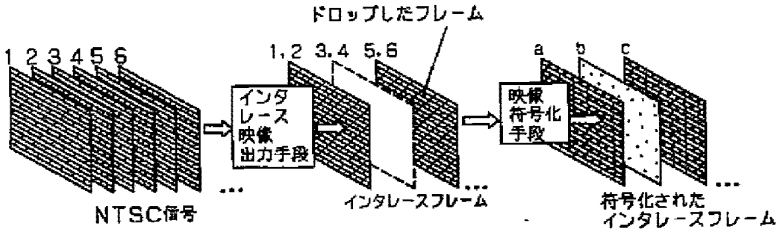
【図１３】


奇数フィールド（240ライン）
偶数フィールド（240ライン）
1201
インタレース映像出力手段
1202
映像符号化手段
1/60sec 1/60sec 1/60sec
NTSC信号
1/30sec
インタレースフレーム
1/30sec
符号化されたインタレースフレーム

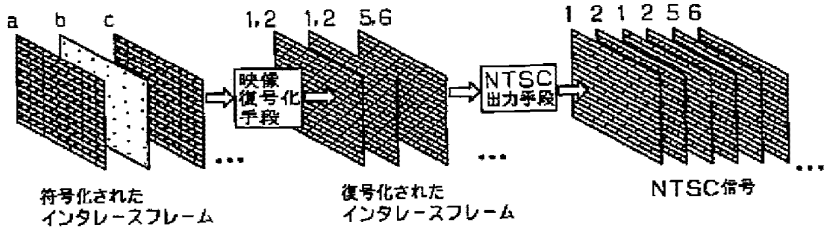
【図１４】



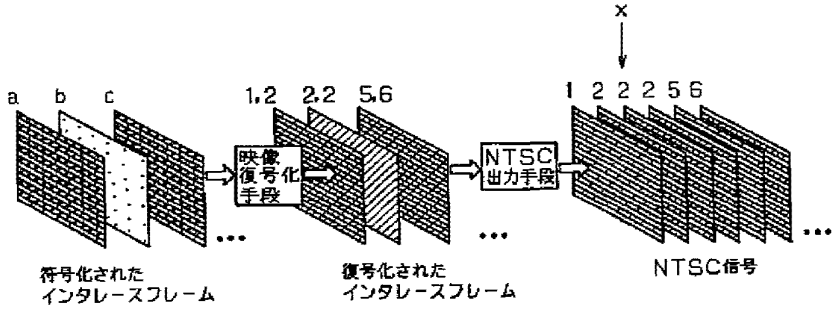
【図１５】



【図１６】



【図１７】



Electronic Acknowledgement Receipt

EFS ID:	6928529
Application Number:	11338934
International Application Number:	
Confirmation Number:	6610
Title of Invention:	Grouping of image frames in video coding
First Named Inventor/Applicant Name:	Miska Hannuksela
Customer Number:	76385
Filer:	Steven R. Funk
Filer Authorized By:	
Attorney Docket Number:	TPO.004.C2
Receipt Date:	08-FEB-2010
Filing Date:	25-JAN-2006
Time Stamp:	10:26:33
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Transmittal Letter	IDS-7-TPO004C2.pdf	55083 df08ecbfb1ca848316b677327a649e6e49a99e4c	no	2

Warnings:	
Information:	

2	Information Disclosure Statement (IDS) Filed (SB/08)	1449-7-TPO004C2.pdf	55900 7c63015565e6ffd90e77f1555c5328f8acfeb6d0	no	1
Warnings:					
Information:					
This is not an USPTO supplied IDS fillable form					
3	Foreign Reference	JP11177987.pdf	816939 84f6ab2c6bf4de5f13d25754972bd8ee4e104810	no	15
Warnings:					
Information:					
4	NPL Documents	20100106-JPOA.pdf	64310 4f9278cb198536d67dee9192adfefa17c1944f59	no	3
Warnings:					
Information:					
		Total Files Size (in bytes):	992232		

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

<u>New Applications Under 35 U.S.C. 111</u>
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

<u>National Stage of an International Application under 35 U.S.C. 371</u>
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

<u>New International Application Filed with the USPTO as a Receiving Office</u>
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant:	HANNUKSELA	Examiner:	Unassigned
Serial No.:	11/338,934	Group Art Unit:	2613
Filed:	January 25, 2006	Docket No.:	TPO.004.C2
Confirmation No.:	6610	Customer No.:	76385
Title:	GROUPING OF IMAGE FRAMES IN VIDEO CODING		

SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT (37 C.F.R. § 1.97(b))

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

With regard to the above-identified application, the items of information listed on the enclosed Form 1449 are brought to the attention of the Examiner.

This statement should be considered because it is submitted before the mailing date of a first Office Action on the merits for the above-identified application. Accordingly, no fee is due for consideration of the items listed on the enclosed Form 1449.

The Japanese reference was first cited in a communication from a foreign patent office regarding a counterpart foreign application not more than three months prior to the filing of this information disclosure statement.

In accordance with 37 C.F.R. § 1.98(a)(2), only a copy of the foreign document or non-U.S. patent/application listed on the enclosed Form 1449 is provided. Also attached is a copy of the English translation of the Japanese Office Action.

Please note that any notations or markings on the attached documents do not reflect particular relevance, or lack thereof, to the present application, nor were they necessarily made by anyone affiliated with the prosecution of the present application.

No representation is made that a reference is "prior art" within the meaning of 35 U.S.C. §§ 102 and 103 and Applicant reserves the right, pursuant to 37 C.F.R. § 1.131 or otherwise, to establish that the reference(s) are not "prior art." Moreover, Applicant does not represent that a reference has been thoroughly reviewed or that any relevance of any portion of a reference is intended.

Consideration of the items listed is respectfully requested. Pursuant to the provisions of MPEP § 609, it is requested that the Examiner return a copy of the attached Form 1449, marked as being considered and initialed by the Examiner, to the undersigned with the next official communication.

Respectfully submitted,

HOLLINGSWORTH & FUNK, LLC
8500 Normandale Lake Blvd., Suite 320
Minneapolis, MN 55437
952.854.2700

By: /Steven R. Funk/
Steven R. Funk
Reg. No. 37,830

FORM 1449* INFORMATION DISCLOSURE STATEMENT IN AN APPLICATION (Use several sheets if necessary)	Docket Number: TPO.004.C2	Application Number: 11/338,934
	Applicant: HANNUKSELA	
	Filing Date: 1/25/2006	Group Art Unit: 2613

U.S. PATENT DOCUMENTS

EXAMINER INITIAL	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	6496980	12/17/2002	Tillman et al.			
	6510553	01/21/2003	Hazra			
	6639943	10/28/2003	Radha et al.			
	6614936	09/02/2003	Wu et al.			

FOREIGN PATENT DOCUMENTS

	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	TRANSLATION	
						YES	NO
	EP0898427	2/24/1999	EP			NA	

OTHER DOCUMENTS (Including Author, Title, Date, Pertinent Pages, Etc.)

	5/7/1996	Elloumi et al., "Issues in Multimedia Scaling: the MPEG Video Streams Case", Emerging Technologies and Applications in Communications, IEEE Computer Society, May 7, 1996.
	2/1/2002	Hannuksela, "Enhanced Concept of GOP", Joint Video Team of ISO-IES MPEG 7 itu-T VCEG, 2/1/2002.
	10/12/1997	Soderquist et al., "Memory Traffic and Data Cache Behavior of an MPEG-2 Software Decoder", IEEE Computer Society, October 12, 1997.
	2/2/2010	Copy of Office Action of related European application no. 06101013.8 dated 02/02/2010.
	2/2/2010	Copy of Office Action of related European application no. 06101014.6 dated 02/02/2010.
	3/16/2006	Copy of Office Action of parallel U.S. Application No. 10/306,942 dated 3/16/2006, 4 pages.
	8/10/2006	Copy of Office Action of parallel U.S. Application No. 10/306,942 dated 8/10/2006, 16 pages.
	9/11/2007	Copy of Office Action of parallel U.S. Application No. 10/306,942 dated 9/11/2007, 4 pages.
	3/13/2007	Copy of Office Action of parallel U.S. Application No. 10/306,942 dated 3/13/2007, 7 pages.
	12/7/2009	Copy of Notice of Allowance of parallel U.S. Application No. 10/306,942 dated 12/7/2009, 8 pages.

EXAMINER	DATE CONSIDERED
EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609; draw line through citation if not in conformance and not considered. Include copy of this form for next communication to the Applicant.	

FORM 1449* INFORMATION DISCLOSURE STATEMENT IN AN APPLICATION (Use several sheets if necessary)	Docket Number: TPO.004.C2	Application Number: 11/338,934
	Applicant: HANNUKSELA	
	Filing Date: 1/25/2006	Group Art Unit: 2613

	5/6/2006	Copy of Office Action Response of parallel U.S. Application No. 10/306,942 dated 8/6/2006, 13 pages.
	11/9/2006	Copy of Office Action Response of parallel U.S. Application No. 10/306,942 dated 11/9/2006, 16 pages.
	5/31/2007	Copy of Office Action Response of parallel U.S. Application No. 10/306,942 dated 5/31/2007, 12 pages.
	7/9/2009	Copy of Office Action Response of parallel U.S. Application No. 10/306,942 dated 7/9/2009, 12 pages.

EXAMINER	DATE CONSIDERED
EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609; draw line through citation if not in conformance and not considered. Include copy of this form for next communication to the Applicant.	

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 0 898 427 A1

(12) EUROPEAN PATENT APPLICATION

published in accordance with Art. 158(3) EPC

(43) Date of publication:
24.02.1999 Bulletin 1999/08

(21) Application number: **98900693.7**

(22) Date of filing: **22.01.1998**

(51) Int. Cl.[6]: **H04N 7/32**

(86) International application number:
PCT/JP98/00232

(87) International publication number:
WO 98/36576 (20.08.1998 Gazette 1998/33)

(84) Designated Contracting States:
DE FR

(30) Priority: **13.02.1997 JP 28862/97**

(71) Applicant:
MITSUBISHI DENKI KABUSHIKI KAISHA
Tokyo 100 (JP)

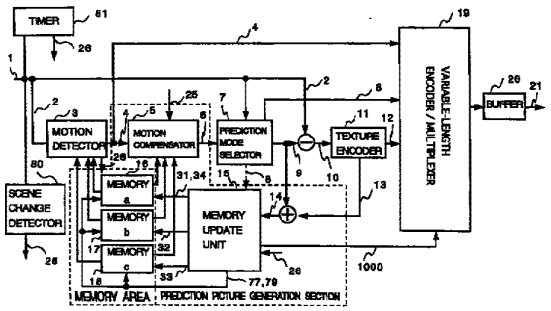
(72) Inventors:
- **SEKIGUCHI, Shunichi,**
  Mitsubishi Denki K.K.
  Chiyoda-ku Tokyo 100 (JP)
- **ASAI, Kohtaro,**
  Mitsubishi Denki K.K.
  Chiyoda-ku Tokyo 100 (JP)
- **MURAKAMI, Tokumichi,**
  Mitsubishi Denki K.K.
  Chiyoda-ku Tokyo 100 (JP)
- **NISHIKAWA, Hirofumi,**
  Mitsubishi Denki K.K.
  Chiyoda-ku Tokyo 100 (JP)
- **KURODA, Shinichi,**
  Mitsubishi Denki K.K.
  Chiyoda-ku Tokyo 100 (JP)
- **ISU, Yoshimi,**
  Mitsubishi Denki K.K.
  Chiyoda-ku Tokyo 100 (JP)
- **HASEGAWA, Yuri,**
  Mitsubishi Denki K.K.
  Chiyoda-ku Tokyo 100 kyo 100 (JP)

(74) Representative:
Pfenning, Meinig & Partner
Mozartstrasse 17
80336 München (DE)

(54) MOVING IMAGE ESTIMATING SYSTEM

(57) To achieve an encoding system including a highly efficient prediction performed in response to the content of a scene, a significance, and a motion characteristic of a moving picture and the like, memories a, b, c, motion compensator 5 responsive to an arbitrary transform parameter representing the motion of a prediction picture segment for generating a predicted picture by using arbitrary data stored in the memories a, b, c based upon the transform parameter, and memory update unit 15 for allowing the content of one or more of the memories to be updated at an arbitrary period of time, are provided.

Fig.1



EP 0 898 427 A1

Description

Technical Field

[0001] The present invention relates to the prediction of a moving picture implemented, for example, in

a moving picture encoder/decoder used in a portable/stationary video communication device and the like for visual communications in a video telephone system, a video conference system or the like,
a moving picture encoder/decoder used in a picture storage/recording apparatus such as a digital VTR and a video server, and
a moving picture encoding/decoding program implemented in the form of a single software or a firmware as a Digital Signal Processor (DSP).

Background Art

[0002] MPEG-4 (Moving Picture Experts Group Phase-4) Video Encoding/Decoding Verification Model (hereinafter referred to by the initials VM) whose standardization is in progress by ISO/IEC JTC1/SC29/WG11 may be introduced as a conventional type of predictive encoding/decoding in an encoding/decoding system of moving pictures. The VM continues to revise its contents according to the progress being made in standardization of MPEG-4. Here, Version 5.0 of the VM is designated to represent the VM and will be simply referred to as VM hereinafter.

[0003] The VM is a system for encoding/decoding each video object as one unit in view of a moving picture sequence being an aggregate of video objects changing their shapes time-/space-wise arbitrarily. Fig. 29 shows a VM video data structure. According to the VM, a time-based moving picture object is called a Video Object (VO), and picture data representing each time instance of the VO, as an encoding unit, is called a Video Object Plane (VOP). If the VO is layered in time/space, a special unit called a Video Object Layer (VOL) is provided between the VO and the VOP for representing a layered VO structure. Each VOP includes shape information and texture information to be separated. If the moving picture sequence includes a single VO, then the VOP is equated to a frame. There is no shape information included, in this case, and the texture information alone is then to be encoded/decoded.

[0004] The VOP includes alpha data representing the shape information and texture data representing the texture information, as illustrated in Fig. 30. Each data are defined as an aggregate of blocks (alphablocks/macroblocks), and each block in the aggregate is composed of 16x16 samples. Each alphablock sample is represented in eight bits. A macroblock includes accompanied chrominance signals being associated with 16x16 sample luminance signals. VOP data are obtained from a moving picture sequence externally processed outside of an encoder.

[0005] Fig. 31 is a diagram showing the configuration of a VOP encoder according to the VM encoding system. The diagram includes original VOP data P1 to be inputted, an alphablock P2 representing the shape information of the VOP, a switch P3a for passing the shape information, if there is any, of the inputted original VOP data, a shape encoder P4 for compressing and encoding the alphablock, compressed alphablock data P5, a locally decoded alphablock P6, texture data (a macroblock) P7, a motion detector P8, a motion parameter P9, a motion compensator P10, a predicted picture candidate P11, a prediction mode selector P12, a prediction mode P13, a predicted picture P14, a prediction error signal P15, a texture encoder P16, texture encoding information P17, a locally decoded prediction error signal P18, a locally decoded macroblock P19, a sprite memory update unit P20, a VOP memory P21, a sprite memory P22, a variable-length encoder/multiplexer P23, a buffer P24, and an encoded bitstream P25.

[0006] Fig. 32 shows a flowchart outlining an operation of the encoder.

[0007] Referring to the encoder of Fig. 31, the original VOP data P1 are decomposed into the alphablocks P2 and the macroblocks P7 (Steps PS2 and PS3). The alphablocks P2 and the macroblocks P7 are transferred to the shape encoder P4 and the motion detector P8, respectively. The shape encoder P4 is a processing block for data compression of the alphablock P2 (step PS4), the process of which is not discussed here further in detail because the compression method of shape information is not particularly relevant to the present invention.

[0008] The shape encoder P4 outputs the compressed alphablock data P5 which is transferred to the variable-length encoder/multiplexer P23, and the locally decoded alpha data P6 which is transferred sequentially to the motion detector P8, the motion compensator P10, the prediction mode selector P12, and the texture encoder P16.

[0009] The motion detector P8, upon reception of the macroblock P7, detects a local-motion vector on a macroblock basis using reference picture data stored in the VOP memory P21 and the locally decoded alphablock P6 (step PS5). Here, the motion vector is one example of a motion parameter. The VOP memory P21 stores the locally decoded picture of a previously encoded VOP. The content of the VOP memory P21 is sequentially updated with the locally decoded picture of a macroblock whenever the macroblock is encoded. In addition, the motion detector P8 detects a global warping parameter, upon reception of the full texture data of the original VOP, by using reference picture data stored in the sprite memory P22 and locally decoded alpha data. The sprite memory P22 will be discussed later in detail.

[0010] The motion compensator P10 generates the predicted picture candidate P11 by using the motion

parameter P9, which is detected in the motion detector P8, and the locally decoded alphablock P6 (step PS6). Then, the prediction mode selector P12 determines the final of the predicted picture P14 and corresponding prediction mode P13 of the macroblock by using a prediction error signal power and an original signal power (step PS7). In addition, the prediction mode selector P12 judges the coding type of the data either intra-frame coding or inter-frame coding.

[0011] The texture encoder P16 processes the prediction error signal P15 or the original macroblock through Discrete Cosine Transformation (DCT) and quantization to obtain a quantized DCT coefficient based upon the prediction mode P13. An obtained quantized DCT coefficient is transferred, directly or after prediction, to the variable-length encoder/multiplexer P23 to be encoded (steps PS8 and PS9). The variable-length encoder/multiplexer P23 converts the received data into a bitstream and multiplexes the data based upon predetermined syntaxes and variable-length codes (step PS10). The quantized DCT coefficient is subject to dequantization and inverse DCT to obtain the locally decoded prediction error signal P18, which is added to the predicted picture P14, and the locally decoded macroblock P19 (step PS11) is obtained. The locally decoded macroblock P19 is written into the VOP memory P21 and the sprite memory P22 to be used for a later VOP prediction (step PS12).

[0012] Dominant portions of prediction including a prediction method, a motion compensation, and the update control of the sprite memory P22 and the VOP memory P21 will be discussed below in detail.

(1) Prediction method in the VM

[0013] Normally, four different types of VOP encoding shown in Fig. 33 are processed in the VM. Each encoding type is associated with a prediction type or method marked by a circle on a macroblock basis. With an I-VOP, intra-frame coding is used singly involving no prediction. With a P-VOP, past VOP data can be used for prediction. With a B-VOP, both past and future VOP data can be used for prediction.

[0014] All the aforementioned prediction types are motion vector based. On the other hand, with a Sprite-VOP, a sprite memory can be used for prediction. The sprite is a picture space generated through a step-by-step mixing process of VOPs based upon a warping parameter set

$$\vec{a}=(a, b, c, d, e, f, g, h)$$

detected on a VOP basis (The mark $\rightarrow$ denotes a vector hereinafter). The warping parameter set is determined by the following parametric equations.

$$x'=(ax+by+c)/(gx+hy+1)$$

$$y'=(dx+ey+f)/(gx+hy+1)$$

The sprite is stored in the sprite memory P22.

[0015] Referring to the parametric equations, (x, y) represents the pixel position of an original VOP in a two-dimensional coordinate system. (x', y') represents a pixel position in the sprite memory corresponding to (x, y,) based upon a warping parameter. With the Sprite-VOP, the warping parameter set can be used uniformly with each macroblock to determine (x', y') in the sprite memory for prediction to generate a predicted picture. In a strict sense, the sprite includes "Dynamic Sprite" used for prediction and "Statistic Sprite" used for prediction as well as for another purpose of an approximate representation of VOP at a decoding station. In Figs. 34 through 37 below, "sprite" stands for Dynamic Sprite.

[0016] The motion detector P8 detects the motion vector and the warping parameter to be used for the aforementioned prediction types. The motion vectors and the warping parameters are generically called the motion parameter P9 hereinafter.

(2) Motion compensation

[0017] Fig. 34 is a diagram showing the configuration of the motion compensator P10 in detail. In the figure, a warping parameter P26, a motion vector P27, a global-motion compensator P28, a local-motion compensator P29, a warping-parameter based predicted picture candidate P30, and a motion-vector based predicted picture candidate P31 are shown. The warping-parameter and motion-vector based predicted picture candidates 30, 31 are generically called the predicted picture candidates P11 hereinafter.

[0018] Fig. 35 shows a flowchart outlining the operation of the motion compensator P10 including steps PS14 through PS21.

[0019] The motion compensator P10 generates the predicted picture candidate P11 using the warping parameter P26 of a full VOP detected on a macroblock P7 basis in the motion detector P8 or a macroblock based motion vector P27. The global-motion compensator P28 performs a motion compensation using the warping parameter P26, and the local-motion compensator P29 performs a motion compensation using the motion vector P27.

[0020] With the I-VOP, the motion compensator P10 does not operate. (The operating step proceeds to step PS21 from step PS14.) With a VOP other than the I-VOP, the local-motion compensator P29 reads out a predicted picture candidate PR1 from the locally decoded picture of a past VOP stored in the VOP memory P21 by using the motion vector P27 (step PS15). With the P-VOP, the predicted picture candidate PR1 is only available to be used.

[0021] When the B-VOP is identified in step PS16, the local-motion compensator P29 further reads out a predicted picture candidate PR2 from the locally decoded

picture of a future VOP stored in the VOP memory P21 by using the motion vector P27 (step PS17). In addition, an arithmetic mean of the predicted picture candidates PR1, PR2 obtained from the past and future VOP locally decoded pictures to obtain a predicted picture candidate PR3 (step PS18).

[0022] A predicted picture candidate PR4 is generated also through Direct Prediction (step PS19). (Direct Prediction is based upon a prediction method corresponding to B-Frame in an encoding method H.263, Recommendation ITU-T. A vector for B-Frame is produced based upon a group of P-VOP vectors, which is not discussed further here in detail.) In Fig. 34, the motion-vector based predicted picture candidates P31 is a generic term for all or part of the predicted picture candidates PR1 through PR4.

[0023] If a VOP is of neither I-VOP nor B-VOP, then the VOP is of Sprite-VOP. With the Sprite-VOP, the predicted picture candidate PR1 is read out from the VOP memory based upon the motion vector. In addition, the global-motion compensator P28 reads out the predicted picture candidate P30 from the sprite memory P22 based upon the warping parameter P26 in step PS20.

[0024] The global-motion compensator P28 calculates the address of a predicted picture candidate in the sprite memory P22 based upon the warping parameter P26, and reads out the predicted picture candidate P30 from the sprite memory P22 to be outputted based upon a resultant address. The local-motion compensator P29 calculates the address of a predicted picture candidate in the VOP memory P21 based upon the motion vector P27 and reads out the predicted picture candidate P31 to be outputted based upon a resultant address.

[0025] These predicted picture candidates P11 are evaluated along with an intra-frame coding signal of the texture data P7 in the prediction mode selector P12, which selects a predicted picture candidate having the least power of a prediction error signal along with a prediction mode.

(3) Updating of memories

[0026] The memory update unit P20 controls the VOP memory P21 and sprite memory P22 to be updated (step PS12). The contents of these memories are updated regardless of the prediction mode P13 selected on a macroblock basis.

[0027] Fig. 36 is a diagram showing the configuration of the memory update unit P20. Fig. 37 shows a flowchart including steps PS22 through PS28 illustrating the operation of the memory update unit P20.

[0028] In Fig. 36, an externally supplied VOP encoding type P32, an externally supplied sprite prediction identification flag P33 for indicating the use of the sprite memory for prediction, an externally supplied blend factor P34 used for prediction with the sprite memory, switches P35, P36, a sprite blender P37, a sprite transformer P38, a VOP memory update signal P39, and a sprite update signal P40 are shown.

[0029] Firstly, the use of the sprite with the current VO or VOL is examined if being designated by the sprite prediction identification flag P33 (step PS22). With no use of the sprite designated, the data are examined if being the B-VOP (step PS27). With the B-VOP, then no updating is performed with the VOP memory P21. With either the I-VOP or the P-VOP, then the VOP memory P21 is written over with the locally decoded macroblock P19 on a macroblock basis (step PS28).

[0030] With the use of the sprite designated in step PS22, then the VOP memory P21 is updated in the same manner as above (steps PS23, PS24), and in addition, the sprite memory PS22 is updated through the following procedure.

a) Sprite warping (step PS25)

[0031] In the sprite transformer P38, an area

$$M(\vec{R},t-1)$$

in the sprite memory P22 ($M(\vec{R},t-1)$ is an area having the same size as that of a VOP having the origin of the coordinates at a position in the sprite memory P22 with the VOP at a time t) is subject to warping (transformation) based upon a warping parameter

$$\vec{\alpha}=(a,b,c,d,e,f,g,h).$$

b) Sprite blending (step PS26)

[0032] By using a resultant warped picture from a) above, a new sprite memory area is calculated in the sprite blender P37 according to the following expression,

$$M(\vec{R},t)=(1-\alpha)\cdot W_b[M(\vec{R},t-1),\vec{\alpha}]+\alpha\cdot VO(\vec{r},t),$$

where α is the blend factor P34, $W_b[M,\vec{\alpha}]$ is the resultant warped picture, and $VO(\vec{r},t)$ is a pixel value of a locally decoded VOP with a location $\vec{r}$ and a time t.

[0033] With a non-VOP area in a locally decoded macroblock, it is assumed that

$$VO(\vec{r},t)=0.$$

As the blend factor α is assigned on a VOP basis, a locally decoded VOP is collectively blended into the sprite memory P22 based upon a weight α, regardless of the contents of a VOP area.

[0034] According to the aforementioned prediction system in the conventional encoding system, the video object is predicted by using the memory designed to be used for detecting the motion vector alone and the memory designed to be used for detecting the warping parameter alone, both of which are structurally allowed the maximum use of a single screen alone each. Thus,

the limited use of reference pictures is only available for prediction, thereby hindering a sufficient improvement in prediction efficiency.

[0035] Further, in such a system where two or more video objects are encoded concurrently, these memories only include a reference picture representing the past record of a video object to be predicted alone, which limits the variation of a reference picture and precludes the utilization of a correlation among video objects for prediction.

[0036] Further, the memories are updated regardless of such items as the internal structure, a characteristic, and the past record of the video object. This results in the insufficient storage of information lacking significant data for predicting a video object, thereby posing a problem of failing to enhance prediction efficiency.

[0037] The present invention is directed to solving the aforementioned problems. An objective of this invention is to provide the prediction system for encoding/decoding of picture data where two or more memories are provided to store the past record of the moving picture sequence effectively in consideration of the internal structure and characteristic of the moving picture sequence, thereby achieving a highly efficient prediction as well as encoding/decoding. In addition, the prediction system provides a sophisticated inter-video object prediction performing among two or more video objects.

Disclosure of the Invention

[0038] According to the present invention, a moving picture prediction system, for predicting a moving picture to be implemented in at least one of an encoder and a decoder, includes a plurality of memories for storing picture data for reference to be used for prediction, the plurality of memories being corresponding to different transform methods, respectively, and a prediction picture generation section for receiving a parameter representing a motion of a picture segment to be predicted, and for generating a predicted picture using the picture data stored in one of the plurality of memories used for the picture segment to be predicted based upon the parameter and one of the transform methods corresponding to the one of the plurality of memories.

[0039] The encoder generates a prediction memory indication information signal indicating the one of the plurality of memories used for generating the predicted picture and transmits the prediction memory indication information signal and the parameter to a decoding station so as to generate the predicted picture using the picture data stored in the one of the plurality of memories based upon the one of the transform methods corresponding to the one of the plurality of memories in the decoding station.

[0040] The decoder receives the parameter and a prediction memory indication information sinal indicating the one of the plurality of memories used for generating the predicted picture from an encoding station, wherein the prediction picture generation section generates the predicted picture using the picture data stored in the one of the plurality of memories based upon the parameter and the one of the transform methods corresponding to the one of the plurality of memories.

[0041] Further, according to the present invention, a moving picture prediction system, for predicting a moving picture to be implemented in at least one of an encoding and a decoding, includes a plurality of memories for storing picture data for reference to be used for prediction, the plurality of memories being assigned to different parameter effective value ranges, respectively, and a prediction picture generation section for receiving a parameter representing a motion of a picture segment to be predicted, for selecting one of the plurality of memories assigned to one of the parameter effective value ranges including a value of the parameter, and for generating a predicted picture using the picture data stored in a selected memory.

[0042] Still further, according to the present invention, a moving picture prediction system, for predicting a moving picture to be implemented in at least one of an encoding and a decoding, includes a plurality of memories for storing picture data for reference to be used for prediction and a prediction picture generation section including a motion compensator for receiving a parameter representing a motion of a picture segment to be predicted, and for generating a predicted picture by using the picture data stored in the plurality of memories based upon the parameter, and a memory update unit for updating the picture data stored in at least one of the plurality of memories at an arbitrary timing.

[0043] The moving picture prediction system predicts the moving picture in a moving picture sequence having first and second video objects, wherein the plurality of memories includes separate first and second pluralities of memories corresponding to the first and second video objects, respectively, and the prediction picture generation section includes separate first and second generators, respectively, corresponding to the first and second video objects, wherein the first generator uses the picture data stored in at least one of the first and second pluralities of memories to generate the predicted picture when predicting the first object, and generates information indicating a use of the second plurality of memories for predicting the first object, the information being added to the predicted picture.

[0044] The prediction picture generation section generates the predicted picture through a change of either one of a number and a size of the plurality of memories in response to a change in the moving picture at each time instance.

[0045] The prediction picture generation section generates the predicted picture in a limited use of memories for prediction in response to a change in the moving picture at each time instance.

[0046] The prediction picture generation section generates the predicted picture by calculating a plurality of

the predicted pictures generated by using the respective picture data stored in the plurality of memories.

[0047] The moving picture prediction system further includes a significance detector for detecting a feature parameter representing a significance of the picture segment to be predicted, wherein the prediction picture generation section generates the predicted picture by selecting at least one of choices of at least one of a plurality of prediction methods, the plurality of memories, and a plurality of memory update methods.

[0048] The moving picture prediction system further includes a significance detector for detecting a parameter representing at least one of an amount of bits available for coding the picture segment to be predicted, an amount of change of the picture segment at each time instance, and a significance of the picture segment, wherein the prediction picture generation section generates the predicted picture by selecting at least one of choices of at least one of a plurality of prediction methods, the plurality of memories, a plurality of memory update methods.

[0049] The moving picture prediction system predicts the moving picture on a video object basis, wherein the moving picture prediction system further includes a significance detector for detecting a parameter representing at least one of an amount of bits available for coding a video object to be predicted, an amount of change in the video object at each time instance, and a significance of the video object, wherein the prediction picture generation section generates the predicted picture by selecting at least one of choices of at least one of a plurality of prediction methods, the plurality of memories, and a plurality of memory update methods.

[0050] The moving picture prediction system further includes a prediction information encoder for encoding prediction relating information of the moving picture, wherein the prediction picture generation section counts times of a memory used for prediction and determines a rank of the plurality of memories based upon a counted number of the times, wherein the prediction information encoder allocates a code length to the prediction relating information to be encoded based upon the rank of a memory used for prediction.

[0051] The plurality of memories includes at least a frame memory for storing the picture data on a frame basis and a sprite memory for storing a sprite picture.

[0052] The sprite memory includes at least one of a dynamic sprite memory involving a regular updating, and a static sprite memory not involving the regular updating.

[0053] The one of the transform methods corresponding to the one of the plurality of memories is at least one of a parallel translation, an affine transformation, and a perspective transformation in an interchangeable manner.

[0054] Still further, according to the present invention, a method for predicting a moving picture to be implemented in at least one of an encoding or a decoding, includes the steps of storing picture data for reference to be used for prediction in a plurality of memories, corresponding different transform methods with the plurality of memories, respectively, receiving a parameter representing a motion of a picture segment to be predicted, and generating a predicted picture using the picture data stored in one of the plurality of memories used for predicting the picture segment based upon the parameter and one of the transform methods being corresponding to the one of the plurality of memories.

[0055] The method for predicting a moving picture further includes the steps of generating a prediction memory indication information signal indicating the one of the plurality of memories used for the picture segment to be predicted, and transmitting the prediction memory indication information signal and the parameter to a decoding station.

[0056] The method for predicting a moving picture is implemented in the decoding, and further includes the step of receiving a prediction memory indication information signal indicating the one of the plurality of memories used for generating the predicted picture and the parameter representing a motion of the picture segment to be predicted from an encoding station.

[0057] Still further, according to the present invention, a method, for predicting a moving picture to be implemented in at least one of an encoding and a decoding, includes the steps of storing picture data for reference to be used for prediction in a plurality of memories, assigning separate parameter effective value ranges to the plurality of memories, respectively, receiving a parameter representing a motion of a picture segment to be predicted, selecting one of the plurality of memories assigned to one of the parameter effective value ranges including a value of the parameter, and generating a predicted picture using the picture data stored in a selected memory.

[0058] Still further, according to the present invention, a method, for predicting a moving picture to be implemented in at least one of an encoding and a decoding, includes the steps of storing picture data for reference to be used for prediction in a plurality of memories, receiving a parameter representing a motion of a picture segment to be predicted, generating a predicted picture using the picture data stored in the plurality of memories based upon the parameter, and updating the picture data stored in at least one of the plurality of memories at an arbitrary timing.

Brief description of drawings

[0059]

Fig. 1 is a structural diagram of a moving picture encoder according to an embodiment of this invention.

Fig. 2 is a flowchart illustrating an operation of the moving picture encoder according to the embodi-

ment of this invention.

Fig. 3 is a structural diagram illustrating the configuration of a motion compensator of the moving picture encoder of the embodiment of this invention.

Fig. 4 is a flowchart illustrating an operation of the motion compensator.

Fig. 5 is a structural diagram illustrating the structure of a memory update unit of the moving picture encoder of the embodiment of this invention.

Fig. 6 is a flowchart illustrating an operation of the memory update unit.

Fig. 7 is a structural diagram illustrating the configuration of a motion compensator of a moving picture encoder according to another embodiment of this invention.

Fig. 8 is a flowchart illustrating an operation of the motion compensator of Fig. 7.

Fig. 9 is a structural diagram illustrating the configuration of a motion compensator of a moving picture encoder according to another embodiment of this invention.

Fig. 10 is a flowchart illustrating an operation of the motion compensator of Fig. 9.

Fig. 11 is a structural diagram of a moving picture encoder according to another embodiment of this invention.

Fig. 12 is a structural diagram showing the configuration of a motion compensator of the moving picture encoder according to the embodiment of this invention.

Fig. 13 is a flowchart illustrating an operation of the motion compensator of Fig. 12;

Fig. 14 is a structural diagram illustrating the configuration of a memory update unit of a moving picture encoder according to another embodiment of this invention.

Fig. 15 is a flowchart illustrating an operation of the memory update unit of Fig. 14.

Fig. 16 is a structural diagram of a moving picture encoder according to another embodiment of this invention.

Fig. 17 is a structural diagram of a moving picture encoder according to another embodiment of this invention.

Fig. 18 is a structural diagram of a moving picture encoder according to another embodiment of this invention.

Fig. 19 is a structural diagram of a moving picture encoder according to another embodiment of this invention.

Fig. 20 is a diagram showing bit stream 21 according to the first embodiment of this invention.

Fig. 21 is a diagram showing bit stream 21 according to the second embodiment of this invention.

Fig. 22 is a diagram showing bit stream 21 according to the third embodiment of this invention.

Fig. 23 is a diagram showing bit stream 21 according to the sixth embodiment of this invention.

Fig. 24 is a structural diagram of a moving picture decoder according to another embodiment of this invention.

Fig. 25 is a structural diagram illustrating the configuration of a motion compensator of the moving picture decoder according to the embodiment of this invention.

Fig. 26 is a flowchart illustrating an operation of the motion compensator.

Fig. 27 is an exemplary diagram of interpolation.

Fig. 28 is a flowchart illustrating an operation of a memory update unit of the moving picture decoder according to the embodiment of this invention.

Fig. 29 is an exemplary diagram of a video data configuration according to the VM encoding system.

Fig. 30 is an exemplary diagram of a VOP data structure.

Fig. 31 is a structural diagram illustrating the configuration of a VM encoder.

Fig. 32 is a flowchart illustrating an operation of the encoder of Fig. 31.

Fig. 33 is an exemplary diagram of VOP encoded types and corresponding prediction types.

Fig. 34 is a structural diagram illustrating the configuration of a motion compensator of the encoder of Fig. 31.

Fig. 35 is a flowchart illustrating an operation of the motion compensator of Fig. 34.

Fig. 36 is a structural diagram illustrating the configuration of a memory update unit of the encoder of Fig. 31.

Fig. 37 is a flowchart illustrating an operation of the memory update unit of Fig. 36.

Best mode for carrying out the invention

Embodiment 1.

[0060] Fig. 1 is a block diagram showing the configuration of an encoder according to a first embodiment and the following embodiments. The diagram illustrates an input moving picture signal 1, texture data 2, a motion detector 3, a motion parameter 4, a motion compensator 5, a predicted picture candidate 6, a prediction mode selector 7, a prediction mode 8, a predicted picture 9, a prediction error picture 10, a texture encoder 11, a quantized DCT coefficient 12, a locally decoded prediction error picture 13, a locally decoded picture 14, a memory update unit 15, a memory-a 16, a memory-b 17, a memory-c 18, a variable-length encoder/multiplexer 19, a transmission buffer 20, a bitstream 21, a scene-change detector 80, and a timer 81. Particularly, the motion compensator 5 and the memory update unit 15 forms a predicted picture generation section 100 which implemented a prediction system. Memories a, b, c forms a memory area 200. In the figure, portions not mentioned in this embodiment will be discussed in the

following embodiments. Fig. 2 shows a flowchart illustrating an operating flow of the encoder.

[0061] This embodiment is based upon the assumption that a plural number, e.g. three, of memories, are used adaptively according to the significance of an input moving picture based upon such characteristics of motion as an amount of motion and an intensity of color. It is also assumed that the content of an arbitrary memory (area), the memory-a for example, are updated at an arbitrary period of time and a moving picture sequence is received on a frame basis.

(1) Input signal

[0062] As aforementioned, the encoder inputs a frame representing a picture at each time instance of the moving picture sequence and decomposes the frame into the encoding units of macroblocks which are one example of picture segments subject to prediction (step S1).

(2) Adaptive use of memories

[0063] The memories store previously decoded pictures or previously provided fixed pictures. In this embodiment, the three memories are used adaptively according to the significance of a picture segment in a frame as follows.

[0064] The memory-a stores a least significant picture segment (i.e., a background-like picture segment whose motion is static or flat and the texture is flat.)

[0065] The memory-b stores a less significant picture segment (i.e., a picture segment of an object whose motion is relatively small.)

[0066] The memory-c stores a most significant picture segment (i.e., a picture segment of an object whose motion is complicated or drastic.)

[0067] The least significant picture segment to be stored in the memory-a may be a background picture segment in a video conference scene or the like. The least significant picture segment also corresponds to a background segment in a camera-work relating flat motion scene of a full screen including slightly moving objects. With this type of motion, it is efficient to obtain a frame based amount of a motion to substitute for a macroblock based motion, rather than to obtain a macroblock based amount of a motion. Specifically, a transform parameter corresponding to the sprite warping parameter discussed in the conventional art is obtained and the transform parameter of a full frame is then used as the motion parameter of a macroblock in the frame. The motion parameter may be selected from among a simple parallel translation parameter (=a motion vector), an affine motion parameter involving transformation, and a perspective motion parameter involving transformation. Here, a motion vector is one example of the motion parameter.

[0068] The less significant picture segment to be stored in the memory-b may be a picture segment of a moving figure who is not a speaker in a video conference scene or the like. This segment type of the object may be considered less attracting in the scene. The most significant picture segment to be stored in the memory-c may be a segment of an object attracting most attention in the video conference scene such as a speaker.

[0069] A picture segment stored in the memory-b or the memory-c representing a unique type of motion of an object should have a macroblock based unique motion parameter. The motion parameter of this case may be selected from among the simple parallel translation parameter (=a motion vector), the affine motion parameter involving transformation, the perspective motion parameter involving transformation, etc.

(3) Motion detection (step S2)

[0070] The motion detector 3 of this embodiment is designed to detect an arbitrary transform parameter of the respective three memories on a macroblock basis, involving no distinction between a motion vector and a warping parameter in the conventional art. The motion detector 3 is provided with additional functions of a global-motion parameter detection for detecting a frame based transform parameter using the memory-a and a local-motion parameter detection for detecting a macroblock based transform parameter using the memories a through c.

(4) Motion compensation (step S3)

[0071] Fig. 3 shows the configuration of the motion compensator 5 of this embodiment in detail. In the figure, a prediction picture memory address calculator 22, a prediction picture memory address 23, a memory reader 24, and a reference memory indicator signal 25 which is supplied externally are shown. In this embodiment, the reference memory indicator signal 25 indicates the use of the memory a, b, c. Fig. 4 shows a flowchart including steps S11 through S16 illustrating an operation of the motion compensator 5.

[0072] Initially, with an I(Intra)-frame, no motion compensation is performed (step S11). With a frame other than the I-frame, predicted picture candidates are generated based upon the global-motion and local-motion parameters corresponding to the respective memories detected in the motion detector 3 (steps S12 through S15). Specifically, the prediction picture memory address calculator 22 calculates the prediction picture memory address 23 of a predicted picture candidate in a memory identified by the reference memory indicator signal 25 based upon the motion parameter 4. Upon reception of the prediction picture memory address 23, the memory reader 24 reads out the predicted picture candidate 6 from a corresponding memory to be outputted.

[0073] In this embodiment, the global-motion and

local-motion parameters are obtained through the same transform method, thereby allowing the motion compensator 5 of Fig. 3 to be shared by both global-motion and local-motion parameter based approaches of generating a predicted picture. When generating the predicted picture candidate 6 through the global-motion parameter, the memory-a is always used as a reference memory (step S15).

(5) Prediction mode selection (step S4)

[0074] The prediction mode of this embodiment is assumed to include the following.

(a) a mode for using the memory-a,
(b) a mode for using the memory-b,
(c) a mode for using the memory-c, and
(d) a mode for using an intra-frame coding signal.

Similarly to the discussion in the conventional art, the prediction mode selector 7 selects the predicted picture candidate 6 having the least power (amplitude) of a prediction error signal, for example, from among all the predicted picture candidates 6 generated in the motion compensator 5 along with an intra-frame coding signal, and outputs a selected one of the predicted picture candidates 6 as the predicted picture 9 and an corresponding one of the prediction mode 8. The prediction mode 8 includes memory selection information indicating a memory used for predicting the selected predicted picture 9. The prediction mode 8 is transferred to the variable-length encoder/multiplexer 19 to be encoded with an allocated length of code in the bitstream 21 as prediction memory indication information 800.

(6) Memory updating

[0075] The memory update unit 15 controls the memories to be updated. Fig. 5 shows the configuration of the memory update unit 15 of this embodiment in detail. In the figure, an activity 26 used for updating the memory-a (which will be discussed later), a memory-a update judger 27, a reference memory selector 28, switches 29, 30, picture data 31 for updating the memory-a, picture data 32 for updating the memory-b, picture data 33 for updating the memory-c, and global prediction picture data 34 for updating the memory-a are shown. Fig. 6 shows a flow of a memory updating operation.

[0076] The memory updating operation of this embodiment has the following procedure. Upon reception of the locally decoded picture 14, the memory update unit 15 judges the necessity of updating the memory-a with a frame including the locally decoded picture 14 in the memory-a update judger 27 (step S17). The reference memory selector 28 selects a memory used for predicting the locally decoded picture based upon the prediction mode 8 (steps S18, S19). Then, a reference picture stored in a selected memory is updated with one of the picture data 31, 32, 33 for updating the memory-a, the memory-b, the memory-c, respectively, and the global prediction picture data 34 for updating the memory-a of the locally decoded picture 14 based upon the following rule. A memory is assumed to be updated each frame on a prediction unit (macroblock) basis.

(1) Frame based regular updating of the memory-b and the memory-c (steps S20, S21):

[0077] The locally decoded picture 14 is written into either the memory-b or the memory-c used for predicting the picture.

(2) Frame based adaptive updating of the memory-a (steps S22, S23):

[0078] The locally decoded picture 14 is written into the memory-a used for predicting the picture for an arbitrary frame only or at an arbitrary period of time based upon a memory-a update judgement 1000 obtained in step S17.

[0079] The content of memory-a is the past record of a time-unvarying picture segment such as a background picture. This removes the necessity of the regular updating of the content of memory unless a full-screen involved complicated or drastic movement such as a scene change occurs to cause a drastic change in the content of a picture segment.

[0080] As aforementioned, a frame based regular updating is performed with a complicated or drastic area of an object, whereas a longer-term based updating is performed with the content of the memory-a, thereby achieving an effective prediction with a background picture half visible among moving objects.

[0081] Viewed in this light, the memory-a is updated in an arbitrary period of time in this embodiment. Specifically, possible arbitrary update criteria are as follows.

a. A full-screen content is updated all at once with a global-motion parameter indicating a complicated or drastic motion, whereas no updating operation is performed with the parameter indicating a rather static motion.
b. A full-screen content is updated all at once on a predetermined period basis, regardless of a frame based period of time.
c. A full-screen content is updated all at once only with a frame immediately after a scene change detected.

[0082] In this embodiment, data as the arbitrary update criteria are generically called as the activity 26 used for updating memory-a. Initially, the memory-a update judger 27 judges whether to update the content of memory-a based upon the activity 26 (step S17). Specifically, the activity 26 corresponds to a value of the

global-motion parameter detected in the motion detector 3 with the arbitrary update criterion-a, a time stamp of the current frame from the timer 81 with the arbitrary update criterion-b, and a flag indicating a scene change detection outputted from the scene-change detector 80 with the arbitrary update criterion-c.

[0083] When the content of memory-a is judged to be updated, the content of the locally decoded picture 14 is outputted as the global prediction picture data 34 for updating the content of the memory-a (step S23). When no updating is judged with the content of the memory-a, then no updating is performed with the memory-a.

[0084] The memory-a update judgement 1000 of a frame is multiplexed in the bitstream 21 to be transmitted to a decoding station so that the same updating of the memory-a can be performed with the frame in the decoding station.

[0085] Fig. 20 is a diagram illustrating the bitstream 21 of this embodiment.

[0086] Fig. 20 is a conceptual diagram showing how frame data are encoded in sequence to be transmitted. Each frame data is provided with header information at the front as a frame based additional information. The header information has the memory-a update judgement 1000 multiplexed to be transmitted to a decoding station. The header information is followed by the component macroblock data of the frame. The macroblock data include the prediction memory indication information 800 indicating a memory used for predicting the macroblock data. In a counterpart decoder, a memory for predicting a predicted picture is specified based upon the prediction memory indication information 800 of macroblock data to generate the predicted picture.

[0087] Although not shown in the figures, the memory-b update information or/and the memory-c update information may be transmitted to the decoding station along with, or alternatively to, the memory-a update judgement 1000.

[0088] The aforementioned encoder thus provides the adaptive and efficient use of two or more memories in response to the content of a moving picture sequence, thereby enhancing prediction efficiency. Specifically, a moving picture sequence is predicted based upon an arbitrary transform parameter through an adaptive use of two or more memories in response to the content and characteristic of a moving picture sequence. This enables an efficient prediction of a moving picture in response to a local characteristic of a picture by even covering complicated motion. With the enhanced prediction efficiency, the encoder is allowed to reduce an amount of encoded data without deteriorating encoded picture quality. The same prediction system may be employed by the counterpart decoder for decoding a bitstream encoded through the prediction system of this invention.

[0089] This embodiment has thus disclosed the encoder performing on a frame basis. The same effect may be expected with an alternative encoder performing on an arbitrary shaped video object (VOP) basis.

[0090] Further, this embodiment has thus disclosed the encoder performing on a macroblock basis as the macroblock being a picture segment subjected to prediction. The same effect can be expected with an alternative encoder for encoding a picture on such a picture segment basis as an arbitrary shaped picture segment and a variable shaped block including fixed-size block components.

[0091] Further, this embodiment has thus disclosed the global-motion parameter detection using the memory-a. Alternatively, a single use of the local-motion parameter detection is of course applicable involving no global-motion parameter detection. With no global-motion detection, no necessity occurs for transmitting prediction information indicating a global/local prediction as the prediction mode.

[0092] Further, this embodiment may include a special memory for prediction which stores reference picture data previously generated based upon the content of the moving picture sequence. The special memory is not updated during an encoding operation.

[0093] Further, this embodiment has thus disclosed the case that memories a, b, c store a picture segment each time and the memory update unit 15 updates one of the memories a, b, c each time. If two or all of the memories a, b, c share to store a picture in part or fully, then the memory update unit 15 updates the two or all of the memories a, b, c. In the case of the memory-a being a frame memory for storing a frame of reference picture data, the memory-b being a static sprite memory involving the adaptive updating, and the memory-c being a dynamic sprite memory involving the regular updating, the memory update unit 15 does not update the memory-b as the static sprite memory for storing previously fixed reference picture data but updates the memory-a and the memory-c concurrently when the memories store the same reference picture segment. Thus, if a duplicated storage of the reference picture data occurs with the memories a, b, c, then the memory update unit 15 updates a duplicated segment stored in each memory.

[0094] The aforementioned can also be applied to the following embodiments.

[0095] Further, this embodiment has thus disclosed the use of three memories a, b, c, but alternatively, two of the memories may be utilized.

[0096] Further, a counterpart decoder may be provided with the prediction picture generation section 100 including the same components as the motion compensator 5 and the memory update unit 15 discussed in this embodiment. A motion compensator provided in the decoder, having no necessity of generating all the three predicted picture candidates, generates a single predicted picture alone based upon a decoded motion parameter.

Embodiment 2.

[0097] A second embodiment shows an encoder with a single replacement of the motion compensator 5 of the encoder shown in Fig. 1. The configuration and operation of a motion compensator 5a of the second embodiment are now described.

[0098] Fig. 7 shows the configuration of the motion compensator 5a of this embodiment in detail. The figure includes a reference memory determiner 35. Fig. 8 shows a flowchart illustrating a detailed operation of the motion compensator 5a.

[0099] Initially, with the I-frame, no compensation is performed (step S24). With a frame other than the I-frame, the reference memory determiner 35 determines a reference memory based upon a value of the motion parameter 4 (step S25). The reference memory determiner 35 holds effective motion parameter value ranges (which will be discussed later in detail) allocated, respectively, to the memories a, b, c. The reference memory determiner 35 compares the respective effective motion parameter value ranges with the value of the motion parameter 4 to judge which memory is designated by the motion parameter 4 and outputs a reference memory indicator signal 25a for identifying the respective memories a, b, c.

[0100] The effective motion parameter value ranges are effective search ranges allocated to the respective memories for detecting a motion vector, for example. Specifically, if ±15 pixels are assumed to be given for a total search value range, then the memory-a is chosen to be used for prediction in a range of ±0 to 3 pixels, the memory-b is used in a range of ±4 to 8 pixels, and the memory-c is used in a range of ±9 to 15 pixels, for example. Here, the reference memory determiner 35 operates only when the local-motion parameter is used for prediction, because the memory-a is exclusively used as a reference memory when the predicted picture candidate is generated based upon the global-motion parameter. This motion-vector value based approach of identifying a memory to be used for prediction is based upon the assumption that a background picture should include a static motion and a most attractive picture should include a complicated or drastic motion. This motion-vector value based approach of identifying a memory for prediction involves no necessity of encoding the prediction mode to be transmitted.

[0101] Next, the predicted picture candidate 6 is generated based upon the reference memory indicator signal 25a of a selected memory (steps S26 through S30). Specifically, the prediction picture memory address calculator 22 calculates the prediction picture memory address 23 of the predicted picture candidate 6 in an identified memory by the reference memory indicator signal 25a based upon the motion parameter 4. Based upon the prediction picture memory address 23, the memory reader 24 reads out the predicted picture candidate 6 from the memory to be outputted.

[0102] As the global-motion and local-motion parameters of this embodiment are based upon the same transform method, both parameter based approaches can share the motion compensator 5a of Fig. 7 for generating the predicted picture candidate. When reading out the predicted picture candidate 6 based upon the global-motion parameter (step S31), the memory-a is always used as a reference memory.

[0103] The effective motion parameter value ranges may be fixed on a moving picture sequence basis, and alternatively, changed on a frame basis, for example. With the frame based changes, the effective motion parameter value ranges assigned to the respective memories of the frame are multiplexed in a bitstream to be transmitted to a decoding station to perform the same memory selection.

[0104] Fig. 21 is a diagram showing the bitstream 21 of this embodiment.

[0105] The bitstream is provided with header information added at the front on a moving picture sequence basis. The header information includes effective motion parameter value range indication information of the respective memories. By thus designating the effective motion parameter value range indication information at the front of a moving picture sequence, the moving picture sequence is predicted with the fixed effective motion parameter value ranges in a counterpart decoder.

[0106] When varying the effective motion parameter value ranges each frame, the effective motion parameter indication information is to be included in the header information added on a frame basis.

[0107] Thus, the efficient and adaptive use of the memories in response to the magnitude of a local motion of a frame can be provided by the encoder including the motion compensator 5a, thereby enhancing prediction efficiency.

[0108] This embodiment has thus disclosed the encoder performing on a frame basis. The same effect may be expected with an alternative encoder performing on an arbitrary shaped video object (VOP) basis.

[0109] Further, this embodiment has thus disclosed the encoder performing on a macroblock basis. The same effect may be expected with an alternative encoder for encoding a picture on such a picture segment basis as an arbitrary shaped picture segment and a variable shaped block including fixed-size block components.

[0110] Further, this embodiment has thus disclosed the global-motion parameter detection using the memory-a. Alternatively, a single use of the local-motion parameter detection is of course applicable involving no global-motion parameter detection. No global-motion detection involves no necessity for transmitting information indicating a global/local prediction as the prediction mode.

[0111] Further, a counterpart decoder may be provided with the prediction picture generation section 100

including the same component as the motion compensator 5 discussed in this embodiment. In the decoder, a motion compensator only generates a single predicted picture based upon a decoded motion parameter.

Embodiment 3.

[0112] Another embodiment shows an encoder having a single replacement of the motion compensator 5 of the encoder of Fig. 1. The configuration and operation of a motion compensator 5b are now described. A motion detector 3a employed in this embodiment is assumed to output an amount of the parallel translation, the affine parameter, and the perspective parameter as motion parameters 4a.

[0113] Further, the memory-a of this embodiment is assumed to be a frame memory for storing a reference picture frame, the memory-b is assumed to be a static sprite memory, and the memory-c is assumed to be a dynamic sprite memory.

[0114] Fig. 9 shows the configuration of the motion compensator 5b of this embodiment in detail. In the figure, a parallel translation amount 36 (i.e., a motion vector), an affine parameter 37, a perspective parameter 38, a parallel-translation based prediction picture memory address calculator 39, an affine parameter based prediction picture memory address calculator 40, and a perspective parameter based prediction picture memory address calculator 41 are shown. Fig. 10 is a flowchart illustrating the operation of the motion compensator 5b in detail.

[0115] Initially, with the I-frame, no prediction is performed (step S33). With a frame other than the I-frame, the prediction picture memory address calculators 39 through 41 calculate the respective prediction picture memory addresses 23 based upon the respective values of the motion parameters 4a (step S34).

[0116] The memory address calculators 39, 40, 41 calculate addresses based upon picture transform methods assigned, respectively, to the corresponding memories. In this embodiment, the parallel translation is assigned to the memory-a, the affine parameter involving such a simple transformation as a rotation and an expansion/contraction is assigned to the memory-b, and the perspective parameter involving a three-dimensional complicated motion is assigned to the memory-c. These transform methods may be expressed by the following transform expressions.

[Parallel translation]

[0117] Amount of parallel translation (a, b):

$$x'=x+a$$

$$y'=y+b$$

(Affine transform]

[0118] Affine parameter (a,b,c,θ):

$$x'=a(\cos\theta)x+a(\sin\theta)y+b$$

$$y'=a(-\sin\theta)x+a(\cos\theta)y+c$$

[Perspective transform]

[0119] Perspective parameter (a,b,c,d,e,f):

$$x'=(ax+by+c)/(gx+hy+1)$$

$$y'=(dx+ey+f)/(gx+hy+1)$$

[0120] Here, (x, y) in a two-dimensional coordinate system represents a pixel location of an original macroblock. (x', y') represents a pixel location in a memory corresponding to (x, y,) based upon each of the parameters. That is a location in a memory (x', y') is calculated based upon these parameters. Through this mechanism, a memory most suitable for the characteristic of a motion can be chosen to be used for prediction on a macroblock basis. With calculated prediction picture memory addresses 23 based upon the respective motion parameters 36, 37, 38, the memory reader 24 reads out the predicted picture candidates 6 from corresponding memories to be outputted (steps S35 through S39).

[0121] The transform methods assigned to the respective memories of the frame are multiplexed in the bitstream 21 to be transmitted to a decoding station as a motion detection method indication information so that the same motion compensation can be performed in the decoding station.

[0122] Fig. 22 is a diagram showing the bitstream 21 of this embodiment.

[0123] Header information added at the front of a moving picture sequence includes the motion detection method indication information. In the encoder, the transformation types to be used in the respective memories are interchangeable, and thus the motion detection method indication information indicating a memory-transform method relation is to be transmitted to the counterpart decoder as the header information of the moving picture sequence. Thus, transformation types assigned to be used with the respective memories can be identified in the decoder.

[0124] In the decoder, the identified transformation types are dynamically assigned to the respective memories.

[0125] Thus, the efficient and adaptive use of the memories in response to the characteristic of a local motion of a frame is provided by the encoder including the motion compensator 5b, thereby enhancing prediction efficiency.

[0126] This embodiment has thus disclosed the

encoder performing on a frame basis. The same effect may be expected with an alternative encoder performing on an arbitrary shaped video object (VOP) basis.

[0127] Further, this embodiment has thus disclosed the encoder performing on a macroblock basis. The same effect may be expected with an alternative encoder for encoding a picture on such a picture segment basis as an arbitrary shaped picture segment and a variable shaped block including fixed-size block components.

[0128] Further, this embodiment has thus disclosed the global-motion parameter detection using the memory-a. Alternatively, a single use of the local-motion parameter detection is of course applicable involving no global-motion parameter detection. No global-motion detection involves no necessity of transmitting information of a global/local prediction as the prediction mode.

[0129] Further, this embodiment has thus disclosed the use of the memories a, b, and c. Alternatively, the use of memories a and b alone, memories a and c alone, or memories b and c alone, is also applicable.

[0130] Further, a decoder may be provided with the prediction picture generating section 100 including the same component as the motion compensator 5b discussed in this embodiment. A motion compensator in a decoder only generates a single predicted picture based upon a decoded motion parameter.

Embodiment 4.

[0131] Another embodiment shows an encoder which receives a plural number, two for example, of different video objects, having shape information, intermingled in a moving picture sequence for a collective encoding. Fig. 11 shows the configuration of the encoder of this embodiment.

[0132] In the figure, an input picture frame 42, an object separator 43, object data 44a, 44b, shape blocks 45a, 45b, switches 46a, 46b, shape encoders 47a, 47b, compressed shape block data 48a, 48b, locally decoded shape blocks 49a, 49b, texture data (macroblocks) 50a,50b, motion detectors 51a, 51b, motion parameters 52a, 52b, motion compensators 53a, 53b, predicted picture candidates 54a, 54b, prediction mode selectors 55a, 55b, prediction mode information 56a, 56b, predicted pictures 57a, 57b, prediction error signals 58a, 58b, texture encoders 59a, 59b, compressed texture data 60a, 60b, locally decoded prediction error signals 61a, 61b, locally decoded macroblocks 62a, 62b, a memory update unit 63, memory-a 64, memory-b 65, memory-c 66, memory-d 67, memory-e 68, memory-f 69, variable-length encoders 70a, 70b, a multiplexer 71, a buffer 72, a bitstream 73, a memory section 94, an object-A encoder 88a for encoding an object-A, and an object-B encoder 88b for encoding an object-B are shown. The object encoders 88a, 88b are structurally identical to each other with the identical components.

[0133] This encoder inputs the picture frame 42, which is decomposed into the encoding units of objects in the object separator 43. The object separator 43 is assumed to be assigned a processing method arbitrarily.

[0134] The shape information of an object is transferred to the shape encoder 47a, 47b in a form of the shape block 45a, 45b to be encoded, and then transferred to the variable-length encoder 70a, 70b as the compressed shape block data 48a, 48b.

[0135] The motion detector 51a, 51b detects a motion parameter based upon the locally decoded shape block 49a, 49b in the same manner as that of the VM encoding system. A motion parameter can be detected on a macroblock basis by using all the memories a through f.

[0136] As a rule, however, the memories a through c are designed to be used for an object-A to be encoded in the object-A encoder 88a, and the memories d through f are designed to be used for an object-B to be encoded in the object-B encoder 88b.

[0137] Also, as for a motion type, an arbitrary transform parameter is assumed to be detected on a macroblock basis with all the memories in the memory section 94, involving no distinction between the motion vector and the warping parameter.

[0138] The motion compensator 53a, 53b generates all the predicted picture candidates 54a, 54b based upon the respective motion parameters 52a, 52b. Then, in the prediction mode selector 55a, 55b, the predicted picture 57a, 57b is obtained along with the prediction mode information 56a, 56b. The predicted picture 57a, 57b is then differentiated from an original signal or the texture data 50a, 50b to obtain the prediction error signal 58a, 58b, which is encoded in the texture encoder 59a, 59b to be transmitted to the variable-length encoder 70a, 70b. The locally decoded prediction error signal 61a, 61b is added to the predicted picture 57a, 57b to obtain the locally decoded macroblock 62a, 62b to be stored into the memories a through f in accordance with an indication by the memory update unit.

[0139] Object A/B data when encoded in the object-A/B encoder 88a, 88b are multiplexed in the bitstream 73 at the multiplexer 71 to be transmitted via the buffer 72.

[0140] The prediction of this embodiment is discussed below focusing on the motion compensator 53a, 53b playing a primary role in the prediction.

[0141] The motion compensator 53a, 53b of this embodiment generates a predicted picture candidate based upon the motion parameter 52a, 52b detected in the motion detector 51a, 51b. Fig. 12 shows the configuration of the motion compensator 53a in detail. Fig. 13 shows a flowchart illustrating the operation of the motion compensator 53a in the object-A encoder 88a.

[0142] In Fig. 12, an object-B reference judger 74a and an object-B reference indicator flag 75a are shown.

[0143] The motion parameter 52a includes memory information used for detection. A predicted picture can-

didate is generated based upon a parameter value through the prediction picture memory address calculator 22a and a memory reader 24a in the same manner as that stated in the first embodiment (step S44 through step S49). The object-B reference judger 74a judges if the memories assigned to object-B are used for predicting the current macroblock based upon the reference memory information included in the motion parameter 52a (step S43).

[0144] The object-B reference judger 74a outputs a judged result as the object-B reference indicator flag 75a, which is multiplexed in the bitstream 73 to be transmitted to a decoding station so as to be used for deciding whether the object can be reproduced in a single use of the memories a, b, c of self in the decoding station. In order to secure the single use of the memories of self when reproducing the object in the decoding station, a limited use of the memories (a, b, c alone) of self for prediction can be controlled by an externally supplied signal 85a at the time of encoding the object.

[0145] Thus, the efficient and adaptive use of the memories in response to the characteristic of a local motion of a frame is provided by the encoder including the motion compensator 53a, 53b, thereby achieving an efficient prediction.

[0146] This embodiment has thus disclosed the encoder for encoding an object on a macroblock basis. The same effect may be expected with an alternative encoder for encoding a picture on such a picture segment basis as an arbitrary shaped picture segment and a variable shaped block including fixed-size block components.

[0147] Further, a decoder may be provided with the same components as the motion compensator 53a, 53b of this embodiment. A motion compensator 53 of the decoder only generates a single predicted picture based upon a decoded motion parameter. Further, if the decoder is structured so as to acknowledge whether a decoding object can be reproduced by itself by way of decoding a bit corresponding to object reference indicator flag 75a, 75b of the other object in a bitstream, then an error-free secured reproduction of decoded object data can be achieved.

Embodiment 5.

[0148] Another embodiment shows an encoder where the number of memories or the size of a memory can be varied flexibly in response to a change in a video object at each time instance. The encoder of a fifth embodiment modifies the encoder of Fig. 1 with a replacement of the memory update unit 15.

[0149] Fig. 14 shows the configuration of a memory update unit 15a of this embodiment in detail. In the figure, a memory expansion judger 76, a memory expansion indicator signal 77, and a memory contraction judger 78, and a memory contraction indicator signal 79 are shown. Fig. 15 shows an operating flow (step S51 through S63) of the memory update unit 15a.

[0150] A picture substantially different from the past record of a moving picture sequence stored in the memories may occur due to a scene change or the like. This may cause a deterioration of prediction efficiency after the scene change if reference pictures stored in the existing memories are the only available. For such an occasion, the scene-change detector 80 detects a scene change, a frame appearing immediately after the detected scene change is subject to intra-frame coding or the like, and resultant intra-frame coded data are stored additionally in a memory as new reference data, thereby enhancing prediction efficiency thereafter.

[0151] Further, in consideration of the physical limitation of storage capacity, a flexible approach of contracting the portions which are rarely used for prediction of the reference pictures stored in the memories is introduced. Specifically, the frequency in use of memory areas for prediction of the respective memories a, b, c is examined in the memory update unit 15a based upon the prediction mode 8. Consequently, the memory update unit releases a memory area identified low in frequency from an area for use. For example, with a software based implementation of this encoder, limited RAM resources may be used effectively.

[0152] Viewed in this light, the memory update unit 15a of this embodiment is provided with a function of expanding a memory area in response to each time instance of a time-varying moving picture sequence and contracting a memory area including a reference picture rarely used for prediction.

[0153] The memory-a, similarly to the first embodiment, is judged in the memory-a update judger 27 whether to be updated (step S50). When updating the memory-a, the locally decoded picture 14 is written into the memory-a (steps S56, S57). The locally decoded picture 14 is written into the other memories as well in accordance with the prediction mode 8 (step S51 through S55).

[0154] The updating of the contents of the memories involves the judgement of memory expansion/contraction. The memory expansion judger 76 judges whether to expand the size of the memory-a (or the memory-b, or the memory-c) based upon the activity 26 used for updating the memory-a (steps S58 through S60). When a positive judgement is made due to a scene change or the like, the expansion of the memory is indicated by the memory expansion indicator signal 77. The memory contraction judger 78 counts the times of a memory area used for prediction based upon the prediction mode 8. With a memory area counted less than a predetermined number in use for prediction, the contraction of the memory area is indicated by the memory contraction indicator signal 79 (steps S61 through S63).

[0155] Thus, a highly efficient prediction can be achieved in response to each time instance of a time-varying moving picture sequence by the encoder including the memory update unit 15a. In addition, the

dynamic allocation of memory areas required for prediction contributes to the enhancement of prediction efficiency and the effective use of memory resources.

[0156] This embodiment has thus disclosed the encoder performing on a frame basis. The same effect may be expected with an alternative encoder performing on an arbitrary shaped video object (VOP) basis.

[0157] Further, this embodiment has thus disclosed the encoder for encoding a frame on a macroblock basis. The same effect can be expected with an alternative encoder for encoding a picture on such a picture segment basis as an arbitrary shaped picture segment and a variable shaped block including fixed-size block components.

[0158] Further, a counterpart decoder may be provided with the same component as the memory update unit 15a discussed in this embodiment.

Embodiment 6.

[0159] With reference to the respective previous embodiments, the memories to be used for prediction are changed on a macroblock basis. Alternatively, the memories to be used for prediction can be changed on a frame or a video object basis. This eliminates the necessity of encoding memory relating information to be encoded on a frame or a video object basis and memory selection information (which is included in the prediction mode 8) to be encoded on a macroblock basis, thereby achieving an efficient encoding.

[0160] With reference to the encoder of Fig. 1 of the first embodiment, for example, the macroblock based changes of the memories used for prediction create the necessity of transmitting additional information identifying a memory used for prediction on a macroblock basis. According to this embodiment, the changing unit of the memories to be used for prediction is limited to a frame or a video object, thereby eliminating the additional information to be transmitted on a macroblock basis effectively. Fig. 23 shows a difference of the transmission bitstream 21 of this embodiment from the transmission bitstream 21 of Fig. 20 of the first embodiment. The bitstream of Fig. 23 represents a frame based change of the memories to be used for prediction with the prediction memory indication information 800 included in frame based header information. The bitstream of Fig. 23 may be effective, for example, in the case that the picture characteristic of a moving picture sequence changes infrequently including little changes locally on a macroblock level. Further, a decoder may be provided so as to decode the thus encoded bitstream to reproduce a frame or a video object.

Embodiment 7.

[0161] With reference to the previous embodiments, two predicted picture candidates read out from an arbitrary plural number, two (e.g., memories a and b) for example, of the memories are subject to an arithmetic mean to obtain a picture as a member of the predicted picture candidates 6 or as the predicted picture 9. Further, a decoder may be provided so as to decode the thus encoded bitstream to reproduce a frame or a video object.

Embodiment 8.

[0162] With reference to the encoders of the previous embodiments, a previously detected feature parameter representing the spatial complexity, perceptual significance and the like of a picture segment as a prediction unit may be utilized as tools for deciding a prediction mode and for judging the updating of the memories.

[0163] For example, a moving picture is assumed to include a motion too complicated to encode data in an acceptable quality within a given amount of encoding. In this case, significance is examined on a prediction picture segment (e.g., a macroblock, an arbitrary shaped picture segment, an arbitrary shaped block) basis. Consequently, a low quality encoding is assigned to some extent to a less significant segment in order to save some amount of encoding for a more significant segment, thereby improving an overall picture quality. With the encoders of this invention where two or more memories are switched interchangeably at an arbitrary timing to be used for prediction, a more adaptive prediction can be achieved in response to the characteristic of a picture, through detecting a feature parameter representing the significance of a prediction picture segment and then determining the use of the memories dynamically based upon a detected feature parameter. For example, as shown in Fig. 16, a segment-significance detector 95 is provided for detecting the feature parameter on a segment basis to determine the significance of the segment. The segment-significance detector 95 transfers a segment-significance to a prediction mode selector 7a and a quantization parameter based upon the segment-significance to a texture encoder 11a. With a segment judged more significant in the segment-significance detector 95, a most complex mode among two or more prediction modes available is used for prediction. Specifically, reference pictures from the respective memories a, b, c are used to obtain the motion parameters and the predicted pictures, respectively, based upon a complex motion model. In the prediction mode selector 7a, a prediction mode having the highest prediction efficiency is selected from among modes including an arbitrary combination (e.g., an arithmetic mean) of the predicted pictures. At the same time, reference pictures of all the memories used for prediction are updated. The texture encoder 11a performs an encoding using a quantization parameter having a smaller quantization step size. With a less significant segment, a simplified prediction mode (i.e., a parallel translation amount detection using a single memory) is employed for prediction and a quantization parameter having a larger

quantization step size is utilized for encoding, regardless of the amplitude of an obtained prediction error signal, so that an amount of encoding be reduced. Through this control, a less significant segment reduces its picture quality to some extent and a more significant segment maintains its quality through a high-quality prediction, thereby improving an overall quality within a given amount of encoding.

Embodiment 9.

[0164] In an encoder where a moving picture sequence is predicted and encoded by using two or more memories, a parameter representing an amount of encoding available for the moving picture sequence at each time instance, an amount of a change in a scene at a certain time instance (e.g., a scene change detection), or the feature parameter or significance of a prediction picture segment described in the eighth embodiment may be detected previously. The values of these parameters may be used for predicting a picture at a particular time instance in a prediction system. Alternatively, these values may be utilized as judging tools for selecting a reference memory area in a prediction system. A frame-significance detector 96 may be provided for determining the significance on a frame basis as shown in Fig. 17. The frame-significance detector 96 detects, for example, an amount of a change in a motion between the current and the previous frame (e.g., a scene change detection by the scene change detector 80), the appearance of a new object or the disappearance of an object or the like. A final significance of the current frame is determined in consideration of an amount of encoding available for the current frame informed by the transmission buffer 20. Based upon the final significance, a more significant frame may be predicted by using all the prediction methods and reference memory areas available for the maximum possible improvement of prediction efficiency, whereas a less significant frame is predicted in a limited use of the prediction methods and the reference memory areas for a simplified encoding so as to reduce the throughput. An alternative encoder for performing intra-frame coding alone involving no prediction at a scene change may be possible. In addition, a more sophisticated quality control may be achieved with a joint use of the segment-significance detector 95 discussed in the eighth embodiment. Through this control, a less significant frame reduces its quality to some extent and a more significant frame maintains its quality by a high-quality prediction, thereby improving an overall picture quality within a given amount of encoding.

[0165] The idea of this embodiment is also applicable to a software based encoding being associated with unsteady transaction processes and unsteady size of available storage in order to achieve an efficient encoding in the maximum use of available resources. This reduces a throughput with a less significant frame, thereby accelerating a general processing speed.

Embodiment 10.

[0166] In an encoder where a moving picture sequence including two or more video objects is predicted and encoded by using two or more memories, as shown in Fig. 11, a parameter representing a gross amount of encoding available for the sequence, an available amount of encoding of a video object at each time instance, an amount of a change in a video object at each particular time instance (e.g., the appearance/disappearance of an object), a level of significance/attention of a video object in a particular scene, or the feature parameter or significance of a prediction picture segment discussed in the eighth and ninth embodiments may be detected previously. The values of these parameters may be utilized for predicting a video object at each particular time instance. Alternatively, these values may be utilized as judging tools for selecting a reference memory area.

[0167] For example, as shown in Fig. 18, significance detectors 97a through 97n responsive, respectively, to objects 1 through n may be provided for detecting a parameter representing an amount of a change in an object at each time instance, or the appearance/disappearance of an object. In addition, the significance of an object at each time instance is determined in consideration of the occupational proportion of a buffer 72x for storing the encoded data of all the objects and the occupational proportion of virtual buffers 72a through 72n for the respective objects. When a new type of segment appears as a result of an object having another object overlapped in part, for example, then the following control may be applied to this type of segment because this type of segment has a great influence on prediction efficiency thereafter. A higher significance may be assigned to this type of segment to obtain an encoded picture in high quality even without enough space available for storage in the corresponding virtual buffer of the object. Significance detected in the significance detectors 97a through 97n is transferred to object 1 through N encoders 98a through 98n, where a full use of the prediction methods and the reference memory areas available is allowed to a more significant object so as to improve prediction efficiency to the maximum, whereas a limited use of the prediction methods and the reference memory areas is assigned to a less significant object so as to simplify the encoding, thereby reducing the throughput. Further, with an encoder for encoding objects decomposed from a frame through a real time separation, when a considerable amount of a change occurs in the contents of the object due to the appearance of a new object or the disappearance of an existing object, the object may be subject to intra-frame coding alone with no prediction involved. A more sophisticated quality control may be achieved on a prediction segment basis of an object in a joint use of the object 1

through N encoders 98a through 98n and the segment-significance detector 95 discussed in the eighth embodiment. Through this control, a less significant object is reduced in quality to some extent and a more significant object manages to maintain its quality through a sophisticated prediction, thereby improving an overall quality within a given amount of encoding.

Embodiment 11.

[0168] An alternative encoder may be provided with a prediction information encoder 91 for allocating a code (encoding) to prediction relating encoding information (e.g., a reference memory number) as shown in Fig. 19.

[0169] In the encoder where a moving picture sequence or a video object is predicted and encoded by using the memories a, b, c, the memories may be ranked based upon frequency in use for prediction with ranks being updated dynamically during an encoding operation. Consequently, a code allocation is performed to the prediction relating encoding information (e.g., a reference memory number) based upon the ranks of the respective memories used for prediction in the prediction information encoder 91.

[0170] For example, in the encoder of Fig. 19, the memory update unit 15b may be provided with a counter 92 which counts the times of the respective memories a, b, c to be used for prediction, ranks the memories a, b, c based upon counted values, and outputs resultant ranking information 90. This ranking may be performed on a picture (VOP) basis at a particular time instance of a frame or a video object, and alternatively, on a smaller unit basis of a prediction picture segment (e.g., a macroblcok, an arbitrary shaped segment, and an arbitrary shaped block).

[0171] This shows how often the respective memories are used for prediction. A memory in a frequent use for prediction is the most significant for prediction, and thus, a high frequency in use for reference corresponding to a high rank.

[0172] When encoding information on the frequency in use for prediction of the memories on a prediction picture segment basis, a memory in a frequent use for reference (i.e., a high-rank memory) is allocated a short code to enhance encoding efficiency.

[0173] In addition, if the motion parameter detected on a prediction picture segment basis is allocated a code length in response to the rank of a memory used for reference, then a shorter code may be assigned to a motion parameter value generated frequently, thereby achieving an efficient encoding of the prediction information. This may be materialized with an alternative encoder where the prediction information encoder 91 in the variable-length encoder/multiplexer 19 receives the ranks of the respective memories from the counter 92 in the memory update unit 15b and encodes the prediction information using a variable-length code based upon the ranking information 90.

Embodiment 12.

[0174] Fig. 24 shows the configuration of a picture decoder where an encoded digital picture through compression is reproduced through expansion according to another embodiment. In the figure, the encoded bitstream 21, a variable-length decoder 119, the quantized DCT coefficient 12, a quantization orthogonal transform coefficient 12a, a quantization step 12b, a texture decoder 111, a dequantizer 111a, an inverse orthogonal transformer 111b, a decoding adder 190, a decoded picture 101, a display controller 191, the prediction mode 8, memory-b update information 1001, memory-c update information 1002, the motion vector 4 (a motion parameter), the prediction memory indication information 800, an in-screen location 195 of a prediction picture segment, a motion compensator 105, a memory-a 116, a memory-b 117, a memory-c 118, a memory update unit 115, and a predicted picture 106 are shown. The motion compensator 105 and the memory update unit 115 form a prediction picture generation section 100a. The memories a, b, c form a memory area 200a.

[0175] According to this embodiment, the memory-a is assumed to be a frame memory designed to store a frame of picture data, the memory-b is assumed to be a static sprite memory, and the memory-c is assumed to be a dynamic sprite memory. The decoder of this embodiment is assumed to receive the bitstream 21 of Fig. 22. Although not shown in Fig. 22, the memory-b update information 1001 and the memory-c update information 1002 are assumed to be transmitted in the bitstream. The memory-b update information 1001 is assumed to include an update indication for a full updating of the static sprite memory and picture data for the full updating. Similarly, the memory-c update information 1002 is assumed to include an update indication for a full updating of the dynamic sprite memory and picture data for the full updating.

[0176] The operation of the thus configured decoder is described below. The variable-length decoder 119 analyzes the bitstream 21 and decomposes it into separate encoded data. The quantization orthogonal transform coefficient 12a is transferred to the dequantizer 119a to be dequantized by using the quantization step 12b. A dequantized result is subject to inverse orthogonal transformation in the inverse orthogonal transformer 111b to obtain a decoded texture, which is transferred to the decoding adder 190. Orthogonal transformation employed here is the same as that employed in an encoding station such as Discrete Cosine Transformation (DCT).

[0177] The motion compensator 105 inputs the motion vector 4, the prediction memory indication information 800, and information indicating the in-screen location 195 of a prediction picture segment included in the bitstream 21, all of which are decoded in the variable-length decoder 119. The motion compensator 105 reads out a right predicted picture from reference pic-

tures stored in the memories a, b, c based upon the three kinds of information. The in-screen location 195 of a prediction picture segment can be obtained by counting the number of macroblocks, other than from the information included in the bitstream. The process of generating a predicted picture will be discussed in a later section for describing the operation of the motion compensator 105 in detail.

[0178] The decoding adder 190, based upon the information of the prediction mode 8, outputs an output from the inverse orthogonal transformer 111b directly as the decoded picture 101 with a block through intra-frame coding, and, with a block through inter-frame coding, adds an output from the inverse orthogonal transformer 111b to the predicted picture 106 to be outputted as the decoded picture 101. The decoded picture 101 is transferred to the display controller 191 to be outputted to a display device and also transferred to the memories a, b, c to be stored as a reference picture for a later use in decoding. A memory writing operation is controlled by the memory update unit 115 based upon the prediction mode 8.

[0179] A predicted picture generation performed in the motion compensator 105 in the prediction picture generation section 100a is now discussed. According to this embodiment, the prediction method of a picture is determined based upon the prediction memory indication information 800. The decoder of this embodiment generates a predicted picture using a reference picture through predetermined coordinate transformation and interpolation based upon the motion vector 4 and the prediction memory indication information 800. Coordinate transform methods are assigned previously to the respective memories to be used for prediction. For example, the following approaches are possible similar to the picture transform methods described in the third embodiment.

(1) the memory-a used for prediction (with the prediction memory identification information 800 indicating the use of the memory-a)

[0180] The coordinates of each pixel of a prediction segment are translated based upon the motion vector and picture data at a corresponding location in the memory-a is read out as the predicted picture.

(2) the memory-b used for prediction (with the prediction memory identification information 800 indicating the use of the memory-b)

[0181] An affine transform expression is found based upon the motion vector, the coordinates of each pixel of a prediction segment are displaced based upon the transform expression, and picture data at a corresponding location in the memory-c is read out as a predicted picture.

(3) the memory-c used for prediction (with the prediction memory identification information 800 indicating the use of the memory-c)

[0182] A perspective transform expression is found based upon the motion vector, the coordinates of each pixel of a prediction segment are displaced based upon the transform expression, and picture data at a corresponding location in the memory-b is read out as a predicted picture.

[0183] Fig. 25 shows the configuration of the motion compensator 105 in detail. In the figure, a switch 161, a corresponding point determiner 162 for the memory-a, a corresponding point determiner 163 for the memory-b, a corresponding point determiner 164 for the memory-c, a memory read address generator 165, a switch 166, and an interpolator 167 are shown. Fig. 26 is a flowchart illustrating the optation of the motion compensator 105.

[0184] The operation of the motion compensator 105 of this embodiment is described below with reference to Figs. 25 and 26.

1) Determining a corresponding point

[0185] Initially, the corresponding point determiner of a corresponding memory is selected by the switch 161 based upon the prediction memory indication information 800. The vector 4 is then inputted to a selected corresponding point determiner. In this section, a predicted picture location corresponding to each memory is calculated, which is explained below with each memory.

1-1) the memory-a indicated by the prediction memory indication information 800 (step S100)

[0186] A predicted picture location is calculated through parallel translation based upon a motion vector (step S101). Specifically, a predicted picture location (x',y') corresponding to a pixel at a prediction picture segment location (x,y) is determined based upon a motion vector (a, b) according to the following expression.

$$x' = x + a$$

$$y' = y + b$$

A determined predicted picture location is outputted to the memory read address generator 165.

1-2) the memory-b indicated by the prediction memory indication information 800 (step S103)

[0187] An affine transform expression is determined based upon the motion vector 4. Specifically, an affine parameter (a, b, c, θ) of the following expression is determined by using the motion vector of a vertex of a rectangular area enclosing a prediction picture seg-

ment,

$$x' = a(\cos\theta)x + a(\sin\theta)y + b$$

$$y' = a(-\sin\theta)x + a(\cos\theta)y + c$$

thereby obtaining the predicted picture location (x',y') corresponding to a pixel at the location (x,y) of a prediction picture segment to be outputted to the memory read address generator 165 (step S104).

1-3) the memory-c indicated by the prediction memory indication information 800 (step S106)

[0188] A perspective transform expression is determined based upon a motion vector. Specifically, a perspective parameter (a, b, c, d, e, f) of the following expression is determined by using the motion vector of a vertex of a rectangular area enclosing a prediction picture segment,

$$x' = (ax + by + c) / (gx + hy + 1)$$

$$y' = (dx + ey + f) / (gx + hy + 1)$$

thereby obtaining the predicted picture location (x',y') corresponding to a pixel at the location (x,y) of a prediction picture segment to be outputted to the memory read address generator (step S107).

2) Reading out data for generating a predicted picture

[0189] Based upon the predicted picture location (x', y') outputted from a selected corresponding point determiner, the memory read address generator 165 generates a memory address for specifying the location of picture data required for generating a predicted picture in a reference picture stored in a memory, and reads out the data for generating a predicted picture (steps S102, 105, 108).

3) Generating a predicted picture

[0190] Among the component pixels of a predicted picture, with a pixel at an integer pixel location, the data for generating a predicted picture is used directly as a component pixel of a predicted picture, and with a pixel at a real number precision pixel location, the data for generating a predicted picture is subject to interpolation in the interpolator 167 to generate an interpolated pixel value (steps S109, S110, S111). Fig. 26 illustrates an interpolated pixel value generation. In Fig. 26, (i_0, j_p) denotes an integer pixel location, (j^p, j^p) denotes a real number precision pixel location, and w denotes a weight.

4) Updating a memory (a reference picture)

[0191] Fig. 28 shows a flowchart illustrating a control operation of the memory update unit 115. The memory update unit 115 controls an updating of the respective memories on a readout unit basis (e.g., a macroblock) of a predicted picture based upon the prediction mode 8 (or the prediction memory indication information 800). With the memory-a used for prediction (step S112), the contents of the memory-a and the memory-c are updated regularly with the decoded picture 101 (step S113). With the memory-b used for prediction (step S114), the reference picture of the memory-b is not updated on a readout unit basis of a predicted picture because of the memory-b being a static sprite memory, whereas the contents of the memory-a and the memory-c are updated regularly with the decoded picture 101 (step S115). When receiving the update indication by the memory-b update information 1001, then the memory update unit updates a full content of the memory-b with received picture data included in the memory-b update information 1001 (step S116). With the memory-c used for prediction (step S117), the contents of the memory-a and the memory-c are updated regularly by using the decoded picture 101 (step S118). When receiving the update indication by the memory update information, the memory update unit updates the content of the memory-c with received picture data included in the memory-c update information 1002 (step S119).

[0192] The use of the three memories a, b, c of this embodiment may be replaced by the use of two memories thereof, for example, with the memories a and b, that is, a frame memory and a static sprite memory. Alternatively, the memories a and c, that is, a frame memory and a dynamic memory, may be used.

[0193] As aforementioned, according to the decoder of this embodiment, the bitstream 21 encoded through an efficient prediction using the various kinds of motion parameters in response to the motion of a picture can be decoded. In addition, the decoder is applicable to the arbitrary updating approach of the contents of a reference picture at a timing determined in the encoding station, thereby achieving a more adaptive decoding in response to the characteristic of a picture.

[0194] According to this embodiment, if the bitstream includes a prediction error signal encoded through encoding other than orthogonal transform encoding, the same effect may be obtained by replacing a component for decoding a prediction error signal, other than the motion compensator and memory update unit.

[0195] Further, this embodiment may be applied not only to a decoder for decoding data on a fixed-size block basis, e.g., for decoding a normal television signal on a frame basis, but also to a decoder for decoding an arbitrary shaped video object (e.g., a Video Object Plane disclosed in ISO/IEC JTC1/SC29/WG11/N1902) as a unit without limiting a prediction segment to a fixed-size

block.

Industrial Feasibility

[0196] As discussed above, the memory areas provided for storing reference pictures according to this embodiment enables the adaptive use of the memories for storing data based upon the characteristic of the moving picture sequence. In addition, the contents of one or more of the memory areas can be updated at an arbitrary timing, so that the content of a time-unvarying picture, such as a background picture, is controlled to be updated on a longer-term basis, and the contents of a locally changing picture segment is controlled to be updated on a regular or sequential basis. This achieves an efficient prediction by reflecting the past record of the moving picture sequence.

[0197] Further, the transform parameter value ranges are assigned to the memory areas for making the respective memory areas effective, and the memory areas are switched to be used for prediction among them based upon the value of the transform parameter of a prediction picture segment, thereby achieving an efficient prediction in response to the magnitude of a local/global motion of the moving picture sequence. At the same time, the motion parameters to be encoded on a prediction picture segment basis can be encoded efficiently within the effective motion parameter value ranges of the reference memory areas.

[0198] Further, to the respective memory areas, the transform methods becoming effective in the respective memories are assigned, and the memories are switched to be used for prediction among them in response to the type of the transform parameter of a prediction picture segment, thereby achieving an efficient prediction in response to the complexity of a local/global motion of the moving picture sequence. At the same time, the transform method can be selected adaptively in response to the characteristic of a prediction picture segment, thereby achieving an efficient encoding of the motion parameter.

Claims

1. A moving picture prediction system for predicting a moving picture to be implemented in at least one of an encoder and a decoder, the moving picture prediction system comprising:

   a plurality of memories for storing picture data for reference to be used for prediction, the plurality of memories corresponding to different transform methods, respectively; and
   a prediction picture generation section for receiving a parameter representing a motion of a picture segment to be predicted, and for generating a predicted picture using the picture data stored in one of the plurality of memories used for the picture segment to be predicted based upon the parameter and one of the transform methods corresponding to the one of the plurality of memories.

2. The moving picture prediction system of claim 1, wherein the encoder generates a prediction memory indication information signal indicating the one of the plurality of memories used for generating the predicted picture and transmits the prediction memory indication information signal and the parameter to a decoding station so as to generate the predicted picture using the picture data stored in the one of the plurality of memories based upon the one of the transform methods corresponding to the one of the plurality of memories in the decoding station.

3. The moving picture prediction system of claim 1, wherein the decoder receives the parameter and a prediction memory indication information sinal indicating the one of the plurality of memories used for generating the predicted picture from an encoding station,

   wherein the prediction picture generation section generates the predicted picture using the picture data stored in the one of the plurality of memories based upon the parameter and the one of the transform methods corresponding to the one of the plurality of memories.

4. A moving picture prediction system for predicting a moving picture to be implemented in at least one of an encoding and a decoding, the moving picture prediction system comprising;

   a plurality of memories for storing picture data for reference to be used for prediction, the plurality of memories being assigned to different parameter effective value ranges, respectively; and
   a prediction picture generation section for receiving a parameter representing a motion of a picture segment to be predicted, for selecting one of the plurality of memories assigned to one of the parameter effective value ranges including a value of the parameter, and for generating a predicted picture using the picture data stored in a selected memory.

5. A moving picture prediction system for predicting a moving picture to be implemented in at least one of an encoding and a decoding, the moving picture prediction system comprising:

   a plurality of memories for storing picture data for reference to be used for prediction; and
   a prediction picture generation section includ-

ing

a motion compensator for receiving a parameter representing a motion of a picture segment to be predicted, and for generating a predicted picture by using the picture data stored in the plurality of memories based upon the parameter, and
a memory update unit for updating the picture data stored in at least one of the plurality of memories at an arbitrary timing.

6. The moving picture prediction system of claim 1, 4, or 5, wherein the moving picture is predicted in a moving picture sequence having first and second video objects,

wherein the plurality of memories includes separate first and second pluralities of memories corresponding to the first and second video objects, respectively, and the prediction picture generation section includes separate first and second generators, respectively, corresponding to the first and second video objects,

wherein the first generator uses the picture data stored in at least one of the first and second pluralities of memories to generate the predicted picture when predicting the first object, and generates information indicating a use of the second plurality of memories for predicting the first object, the information being added to the predicted picture.

7. The moving picture prediction system of claim 1, 4 or 5, wherein the prediction picture generation section generates the predicted picture through a change of either one of a number and a size of the plurality of memories in response to a change in the moving picture at each time instance.

8. The moving picture prediction system of claim 1, 4 or 5, wherein the prediction picture generation section generates the predicted picture in a limited use of memories for prediction in response to a change in the moving picture at each time instance.

9. The moving picture prediction system of claim 1, wherein the prediction picture generation section generates the predicted picture by calculating a plurality of the predicted pictures generated by using the respective picture data stored in the plurality of memories.

10. The moving picture prediction system of claim 1 further comprising:

a significance detector for detecting a feature parameter representing a significance of the picture segment to be predicted,

wherein the prediction picture generation section generates the predicted picture by selecting at least one of choices of at least one of a plurality of prediction methods, the plurality of memories, and a plurality of memory update methods.

11. The moving picture prediction system of claim 1 or 5 further comprising:

a significance detector for detecting a parameter representing at least one of an amount of bits available for coding the picture segment to be predicted, an amount of change of the picture segment at each time instance, and a significance of the picture segment;

wherein the prediction picture generation section generates the predicted picture by selecting at least one of choices of at least one of a plurality of prediction methods, the plurality of memories, a plurality of memory update methods.

12. The moving picture prediction system of claim 1 or 5, wherein the moving picture is predicted on a video object basis,

wherein the moving picture prediction system further comprising:

a significance detector for detecting a parameter representing at least one of an amount of bits available for coding a video object to be predicted, an amount of change in the video object at each time instance, and a significance of the video object;

wherein the prediction picture generation section generates the predicted picture by selecting at least one of choices of at least one of a plurality of prediction methods, the plurality of memories, and a plurality of memory update methods.

13. The moving picture prediction system of claim 1 or 5 further comprising:

a prediction information encoder for encoding prediction relating information of the moving picture;

wherein the prediction picture generation section counts a number of times of a memory used for prediction and determines a rank of the plurality of memories based upon a counted number of the times,

wherein the prediction information encoder allocates a code length to the prediction relating information to be encoded based upon the rank of a memory used for prediction.

14. The moving picture prediction system of claim 1, wherein the plurality of memories includes at least a frame memory for storing the picture data on a frame basis and a sprite memory for storing a sprite picture.

15. The moving picture prediction system of claim 14, wherein the sprite memory includes at least one of a dynamic sprite memory involving a regular updating, and a static sprite memory not involving the regular updating.

16. The moving picture prediction system of claim 14, wherein the one of the transform methods corresponding to the one of the plurality of memories is at least one of a parallel translation, an affine transformation, and a perspective transformation in an interchangeable manner.

17. A method for predicting a moving picture to be implemented in at least one of an encoding or a decoding, the method comprising the steps of:

storing picture data for reference to be used for prediction in a plurality of memories;
corresponding different transform methods with the plurality of memories, respectively;
receiving a parameter representing a motion of a picture segment to be predicted; and
generating a predicted picture using the picture data stored in one of the plurality of memories used for predicting the picture segment based upon the parameter and one of the transform methods being corresponding to the one of the plurality of memories.

18. The method for predicting a moving picture of claim 17, further comprising the steps of:

generating a prediction memory indication information signal indicating the one of the plurality of memories used for the picture segment to be predicted; and
transmitting the prediction memory indication information signal and the parameter to a decoding station.

19. The method for predicting a moving picture of claim 17 to be implemented in the decoding, the method further comprising the step of:

receiving a prediction memory indication information signal indicating the one of the plurality of memories used for generating the predicted picture and the parameter representing a motion of the picture segment to be predicted from an encoding station.

20. A method for predicting a moving picture to be implemented in at least one of an encoding and a decoding, the method comprising the steps of:

storing picture data for reference to be used for prediction in a plurality of memories;
assigning separate parameter effective value ranges to the plurality of memories, respectively;
receiving a parameter representing a motion of a picture segment to be predicted;
selecting one of the plurality of memories assigned to one of the parameter effective value ranges including a value of the parameter; and
generating a predicted picture using the picture data stored in a selected memory.

21. A method for predicting a moving picture to be implemented in at least one of an encoding and a decoding, the method comprising the steps of:

storing picture data for reference to be used for prediction in a plurality of memories;
receiving a parameter representing a motion of a picture segment to be predicted;
generating a predicted picture using the picture data stored in the plurality of memories based upon the parameter; and
updating the picture data stored in at least one of the plurality of memories at an arbitrary timing.

Fig.1

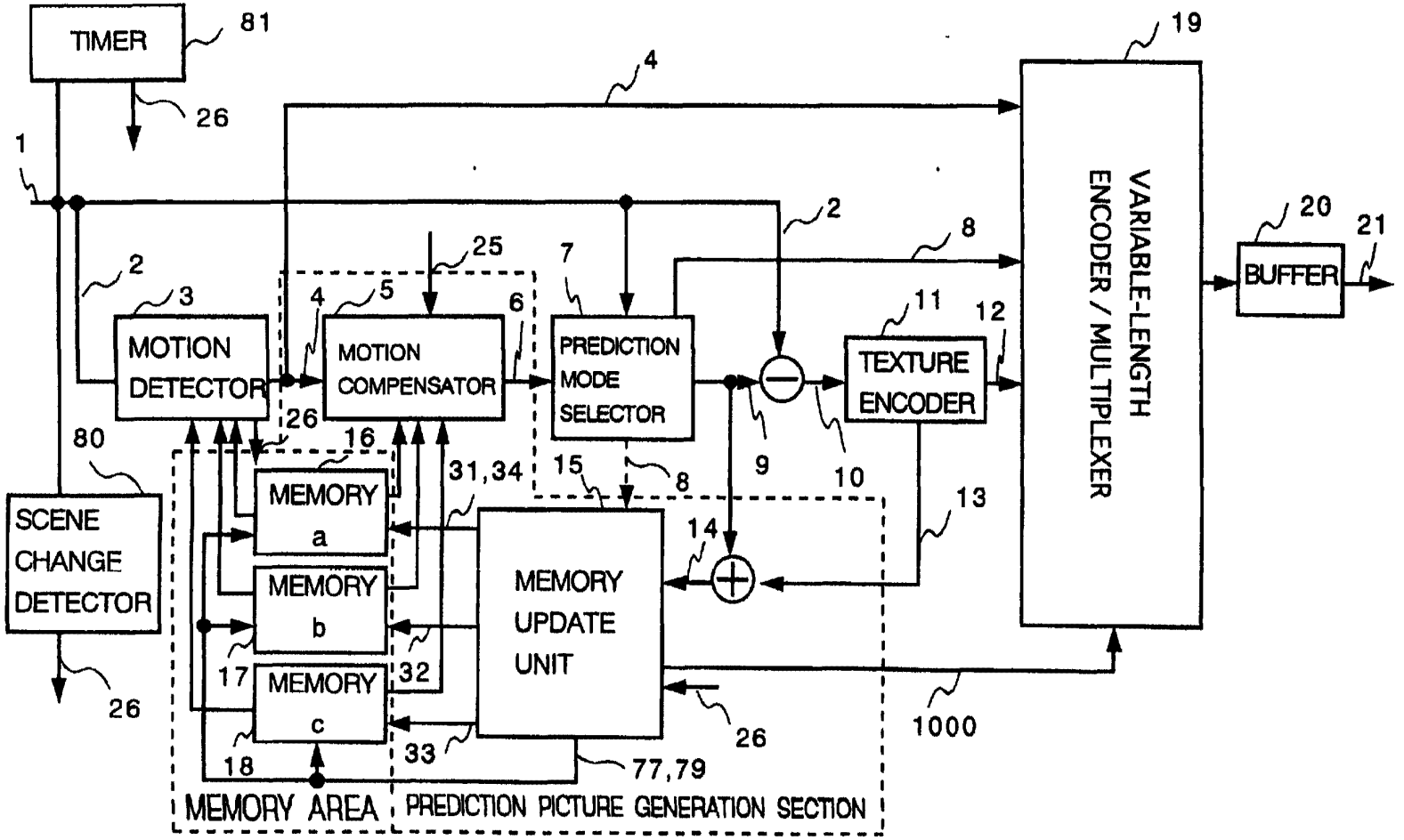

TIMER
81
26
1
4
19
VARIABLE-LENGTH ENCODER / MULTIPLEXER
2
8
20
21
BUFFER
25
7
2
3
4
5
6
11
12
MOTION DETECTOR
MOTION COMPENSATOR
PREDICTION MODE SELECTOR
TEXTURE ENCODER
26
16
80
31,34
15
8
9
10
13
MEMORY a
MEMORY b
MEMORY c
SCENE CHANGE DETECTOR
14
MEMORY UPDATE UNIT
17
32
26
26
1000
18
33
77,79
MEMORY AREA
PREDICTION PICTURE GENERATION SECTION

Fig.2

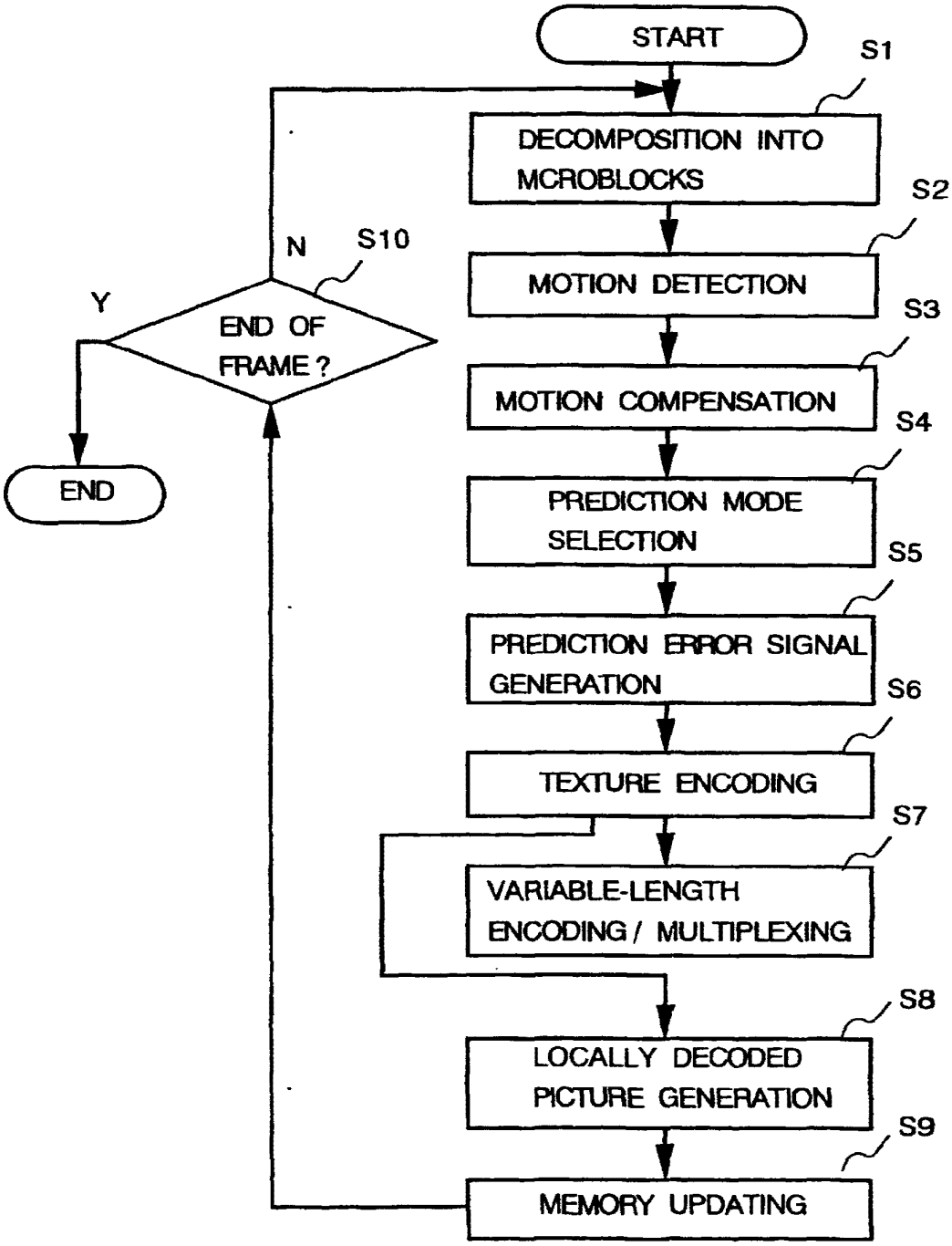

START
S1
DECOMPOSITION INTO MCROBLOCKS
S2
MOTION DETECTION
S3
MOTION COMPENSATION
S4
PREDICTION MODE SELECTION
S5
PREDICTION ERROR SIGNAL GENERATION
S6
TEXTURE ENCODING
S7
VARIABLE-LENGTH ENCODING / MULTIPLEXING
S8
LOCALLY DECODED PICTURE GENERATION
S9
MEMORY UPDATING
S10
END OF FRAME ?
N
Y
END

Fig.3

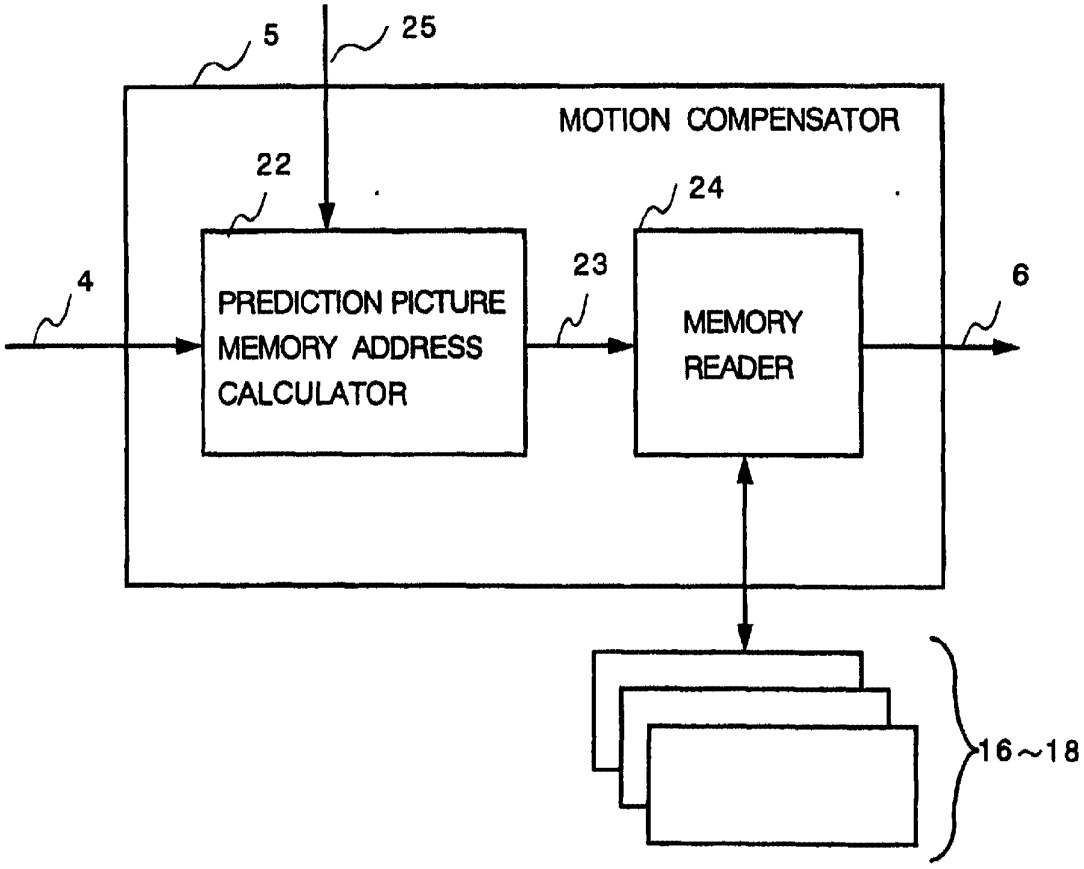

5
25
MOTION COMPENSATOR
22
24
4
PREDICTION PICTURE MEMORY ADDRESS CALCULATOR
23
MEMORY READER
6
16~18

Fig.4

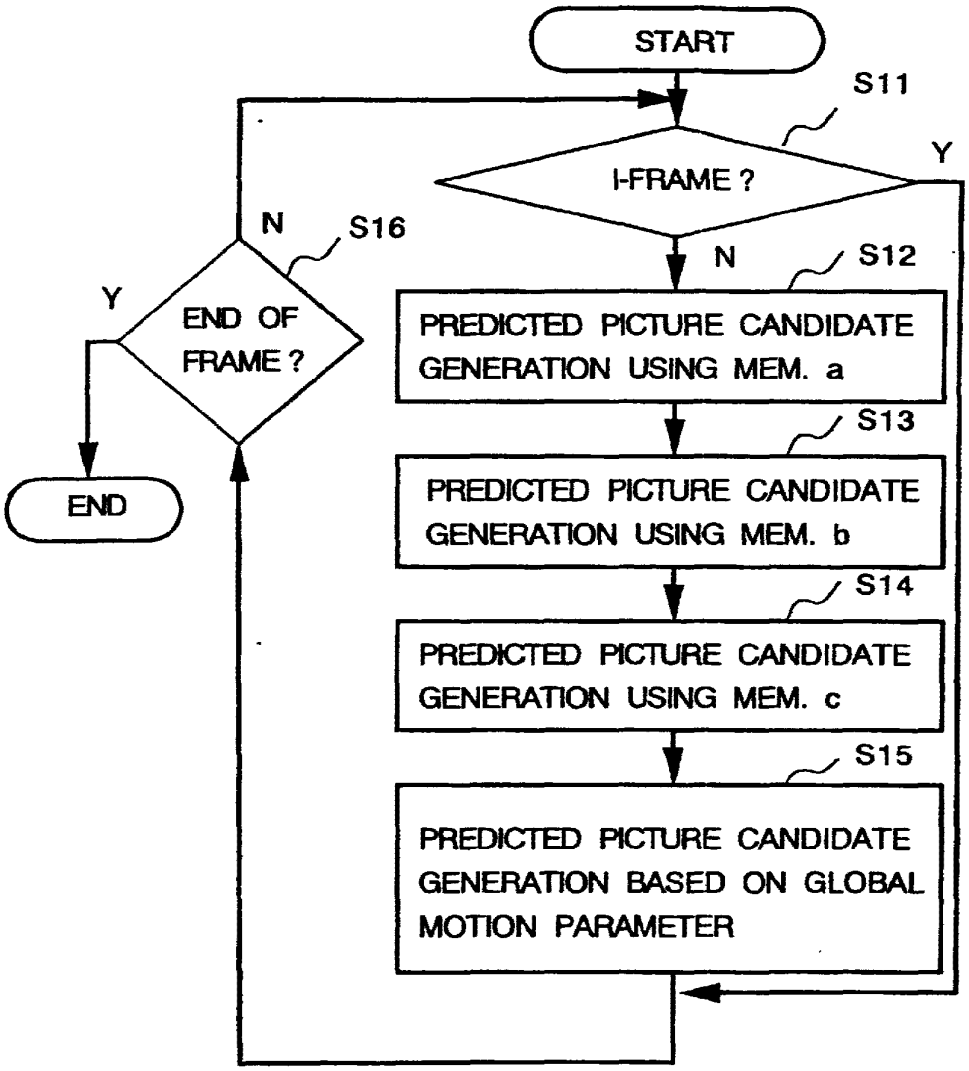

START
S11
I-FRAME ?
Y
N
S12
PREDICTED PICTURE CANDIDATE GENERATION USING MEM. a
S13
PREDICTED PICTURE CANDIDATE GENERATION USING MEM. b
S14
PREDICTED PICTURE CANDIDATE GENERATION USING MEM. c
S15
PREDICTED PICTURE CANDIDATE GENERATION BASED ON GLOBAL MOTION PARAMETER
N
S16
END OF FRAME ?
Y
END

Fig.5

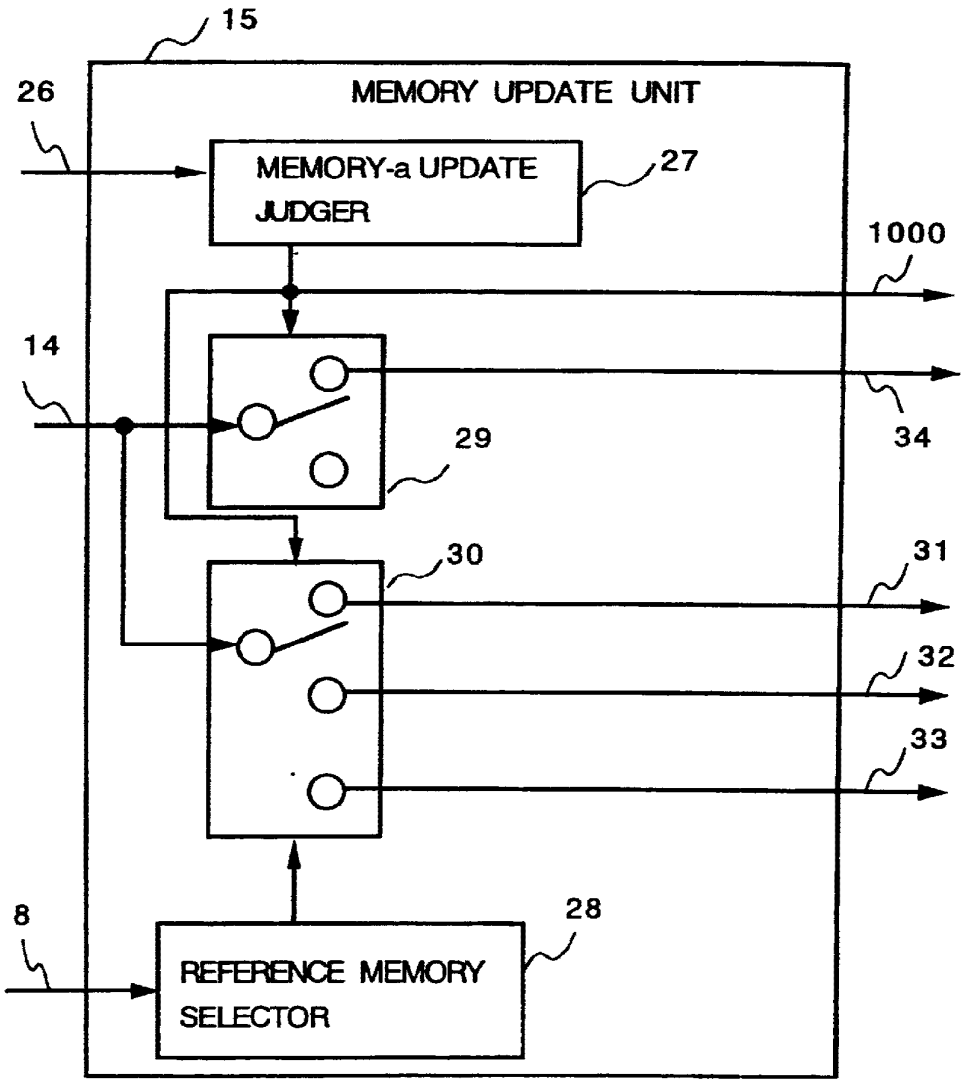

15
MEMORY UPDATE UNIT
26
MEMORY-a UPDATE JUDGER
27
1000
14
29
34
30
31
32
33
8
28
REFERENCE MEMORY SELECTOR

Fig.6

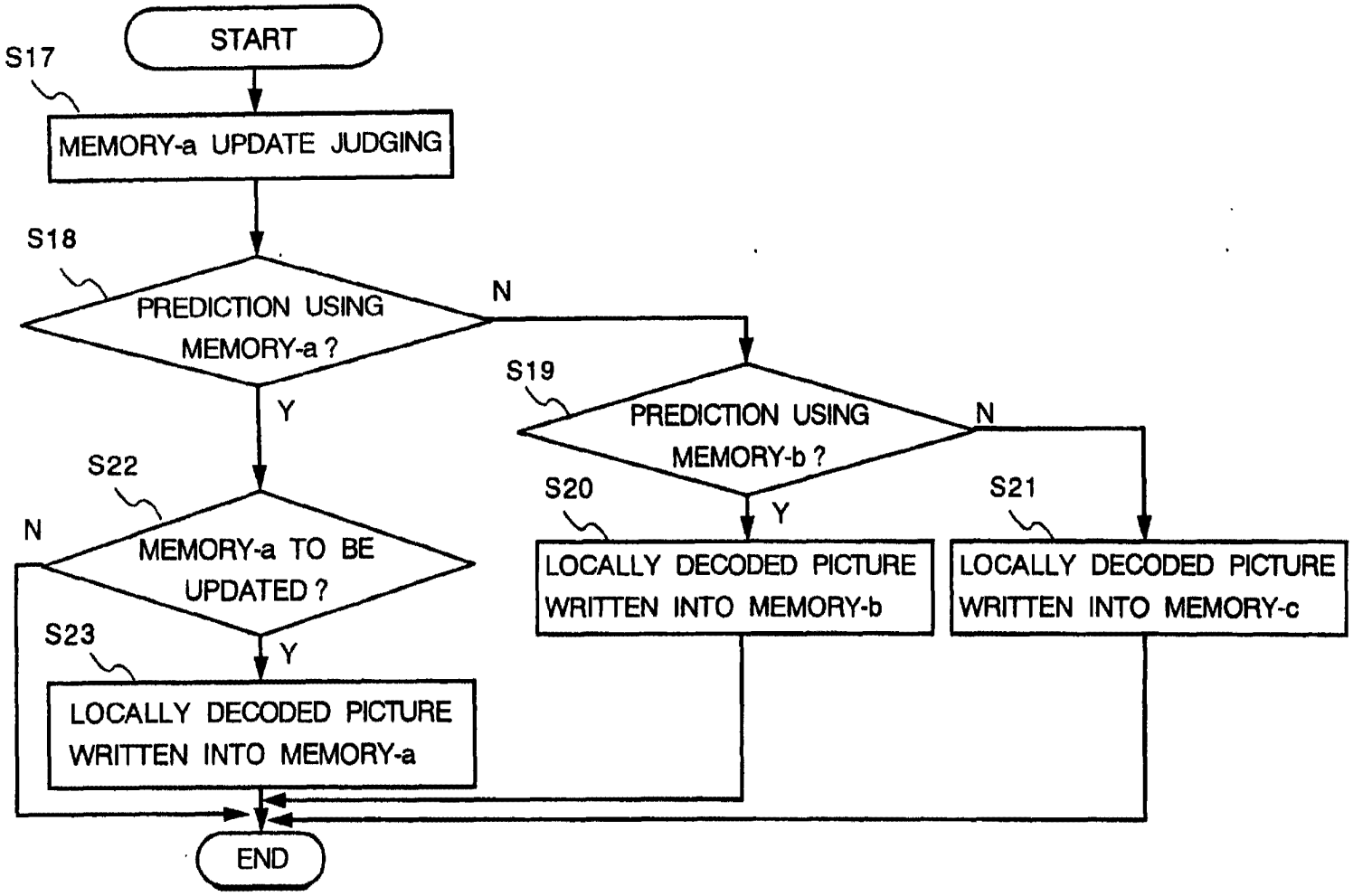

START
S17
MEMORY-a UPDATE JUDGING
S18
PREDICTION USING MEMORY-a ?
N
Y
S19
PREDICTION USING MEMORY-b ?
N
Y
S20
LOCALLY DECODED PICTURE WRITTEN INTO MEMORY-b
S21
LOCALLY DECODED PICTURE WRITTEN INTO MEMORY-c
S22
MEMORY-a TO BE UPDATED ?
N
Y
S23
LOCALLY DECODED PICTURE WRITTEN INTO MEMORY-a
END

Fig.7

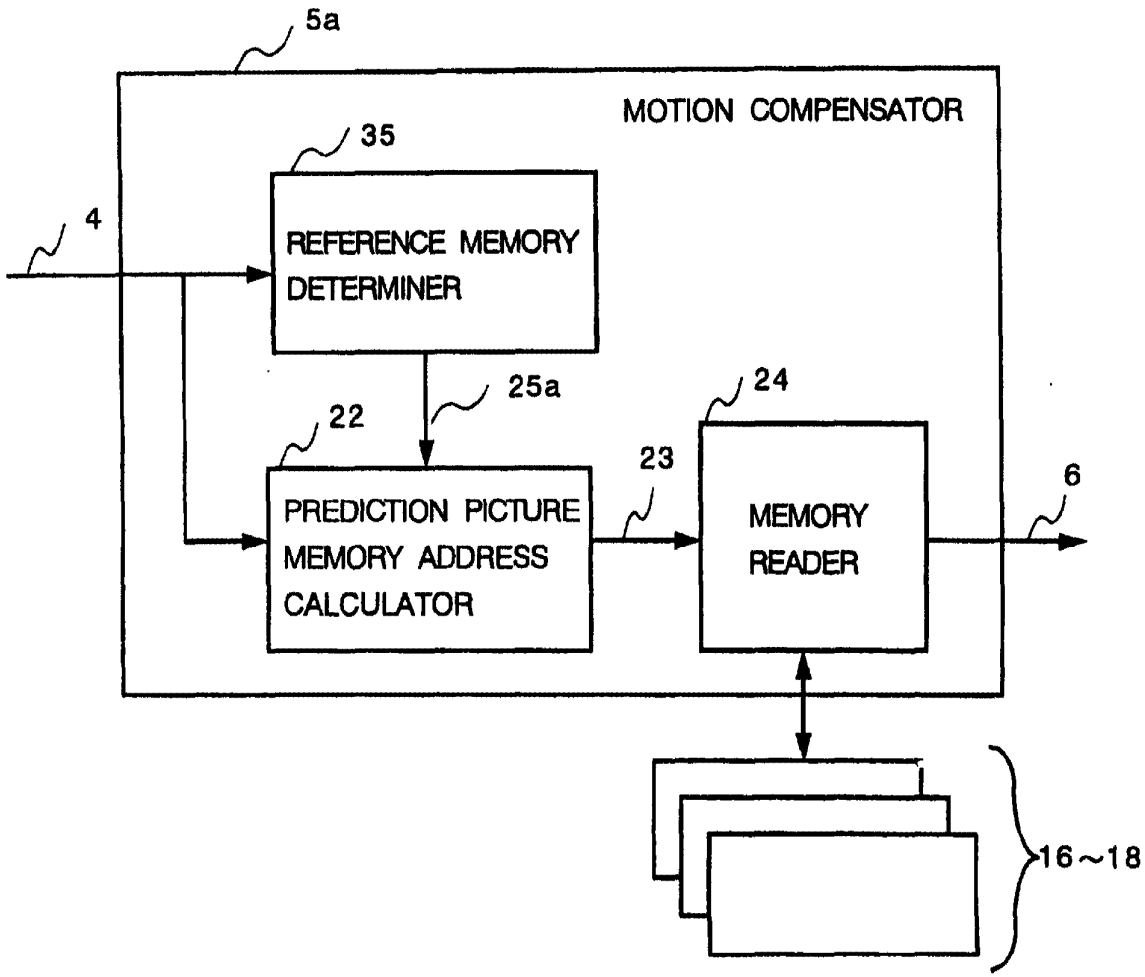

5a
MOTION COMPENSATOR
35
4
REFERENCE MEMORY DETERMINER
25a
24
22
23
6
PREDICTION PICTURE MEMORY ADDRESS CALCULATOR
MEMORY READER
16~18

Fig.8

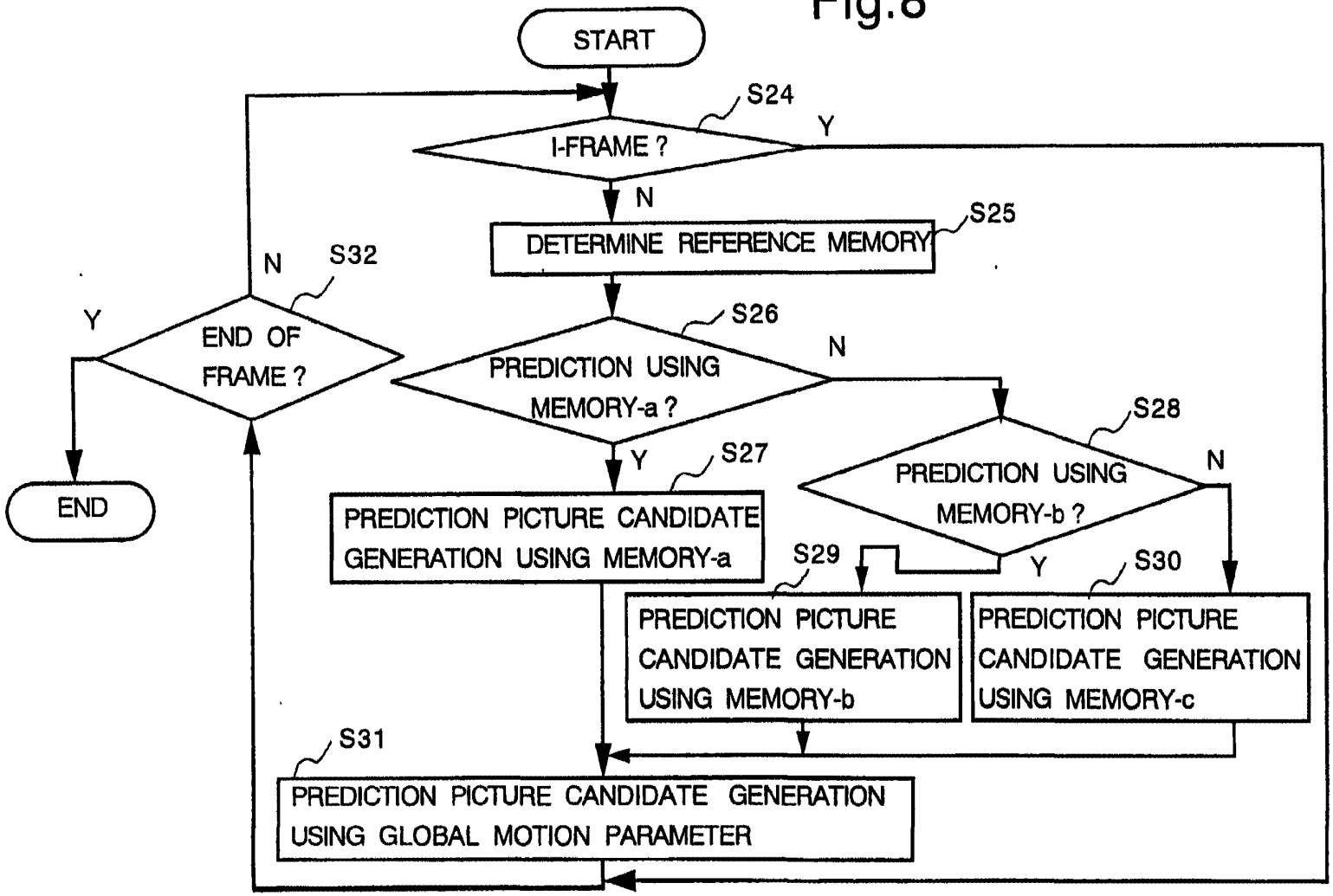

START
S24
I-FRAME ?
Y
N
S25
DETERMINE REFERENCE MEMORY
S26
PREDICTION USING MEMORY-a ?
N
Y
S27
PREDICTION PICTURE CANDIDATE GENERATION USING MEMORY-a
S28
PREDICTION USING MEMORY-b ?
N
Y
S29
PREDICTION PICTURE CANDIDATE GENERATION USING MEMORY-b
S30
PREDICTION PICTURE CANDIDATE GENERATION USING MEMORY-c
S31
PREDICTION PICTURE CANDIDATE GENERATION USING GLOBAL MOTION PARAMETER
S32
END OF FRAME ?
N
Y
END

Fig.9

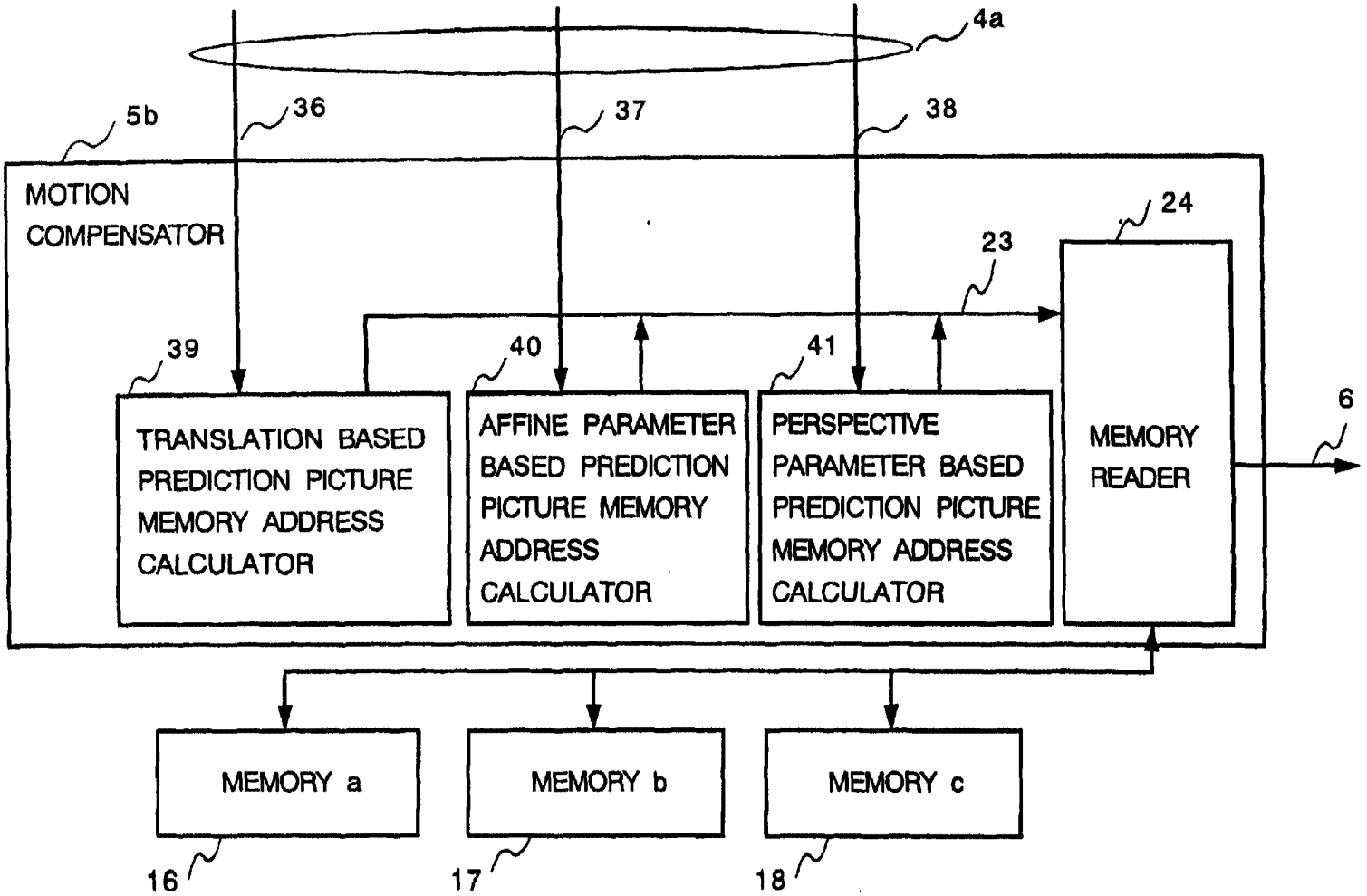

4a
5b
36
37
38
MOTION COMPENSATOR
24
23
39
40
41
TRANSLATION BASED PREDICTION PICTURE MEMORY ADDRESS CALCULATOR
AFFINE PARAMETER BASED PREDICTION PICTURE MEMORY ADDRESS CALCULATOR
PERSPECTIVE PARAMETER BASED PREDICTION PICTURE MEMORY ADDRESS CALCULATOR
MEMORY READER
6
MEMORY a
MEMORY b
MEMORY c
16
17
18

Fig.10

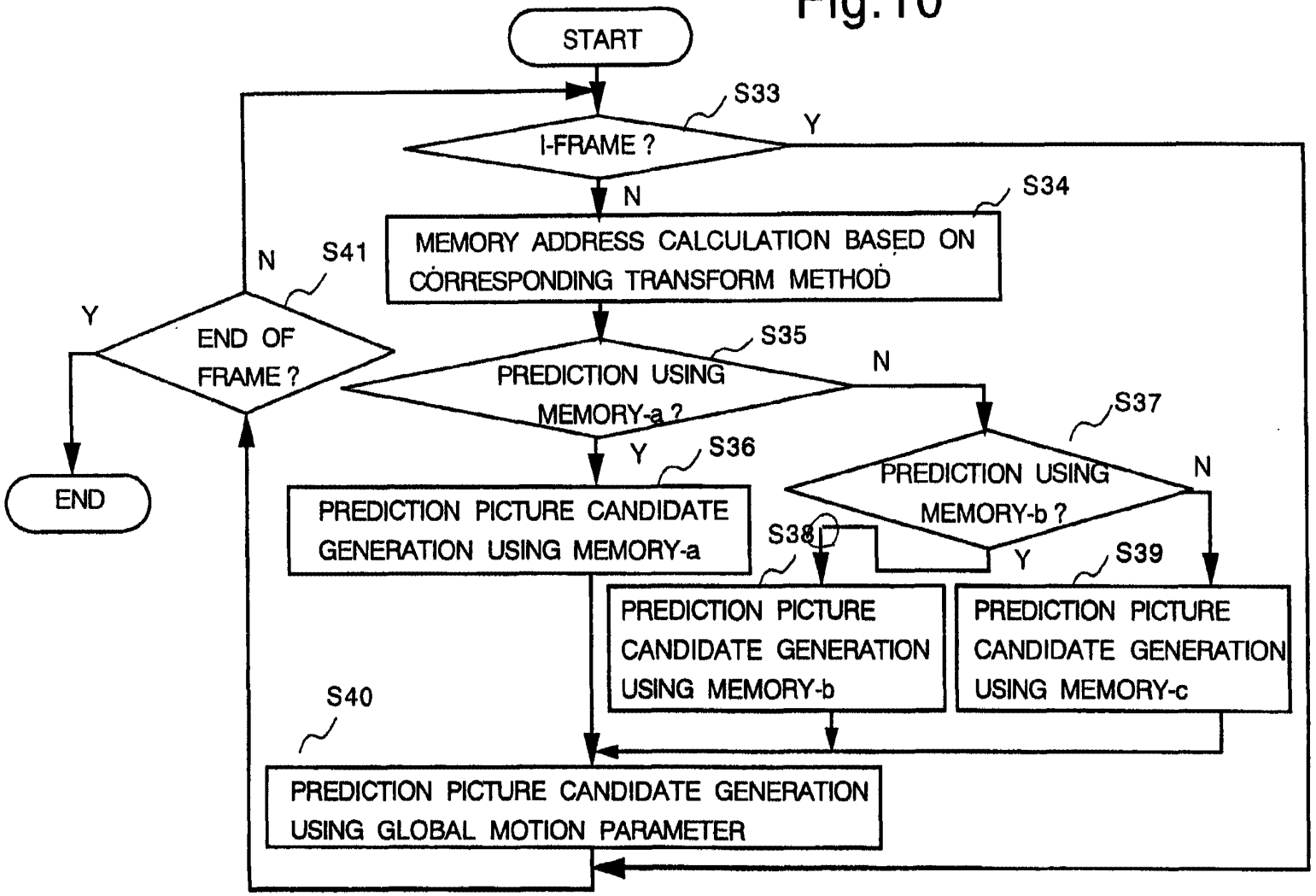

START
S33
I-FRAME ?
Y
N
S34
MEMORY ADDRESS CALCULATION BASED ON CORRESPONDING TRANSFORM METHOD
S35
PREDICTION USING MEMORY-a ?
N
Y
S36
PREDICTION PICTURE CANDIDATE GENERATION USING MEMORY-a
S37
PREDICTION USING MEMORY-b ?
N
Y
S38
S39
PREDICTION PICTURE CANDIDATE GENERATION USING MEMORY-b
PREDICTION PICTURE CANDIDATE GENERATION USING MEMORY-c
S40
PREDICTION PICTURE CANDIDATE GENERATION USING GLOBAL MOTION PARAMETER
S41
N
Y
END OF FRAME ?
END

Fig.11

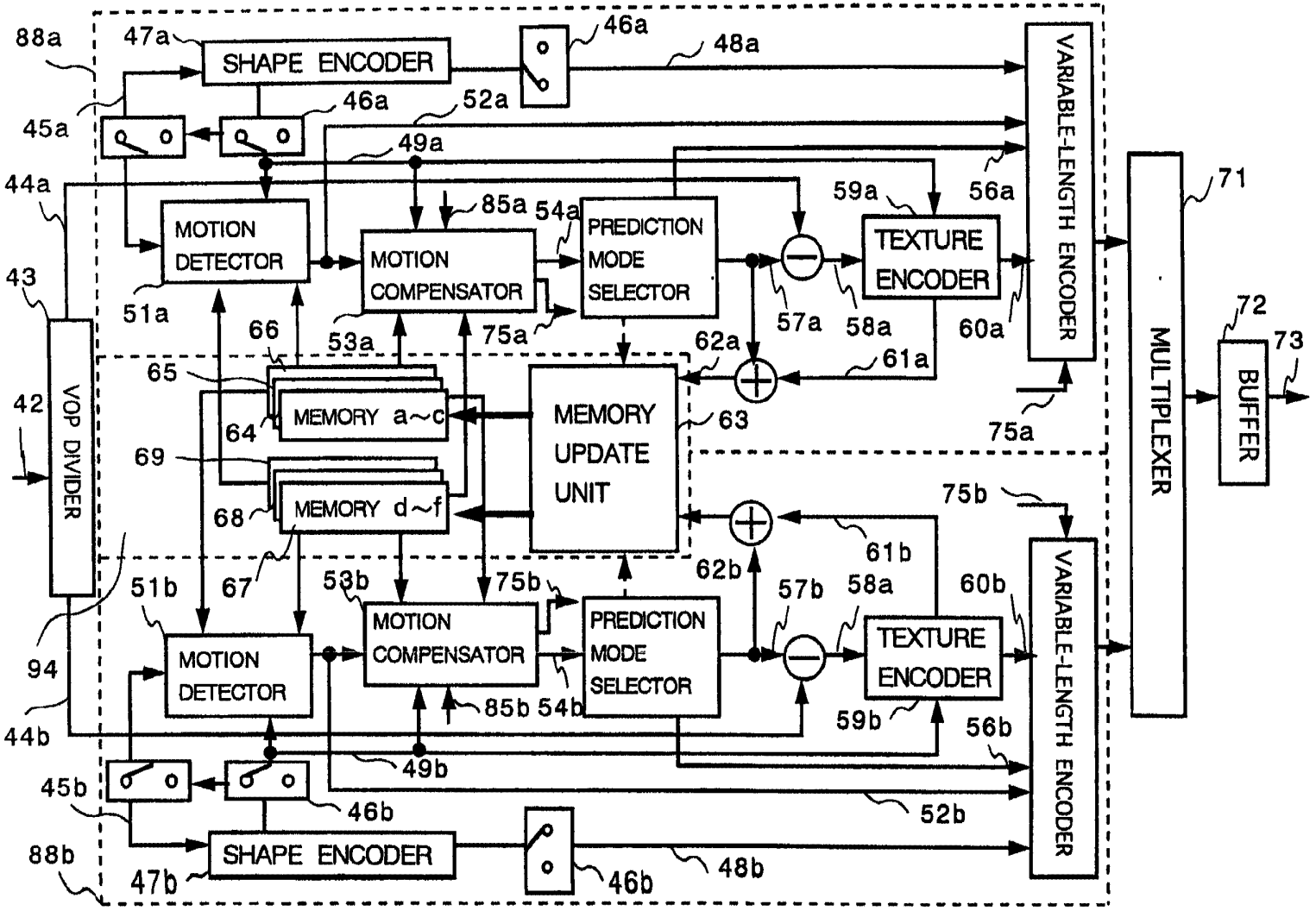

88a
47a
SHAPE ENCODER
46a
48a
45a
46a
52a
49a
44a
MOTION DETECTOR
85a
54a
PREDICTION MODE SELECTOR
59a
56a
TEXTURE ENCODER
VARIABLE-LENGTH ENCODER
71
MULTIPLEXER
72
BUFFER
73
43
51a
66
MOTION COMPENSATOR
53a
75a
57a
58a
60a
62a
61a
65
42
VOP DIVIDER
64
MEMORY a~c
MEMORY UPDATE UNIT
63
75a
69
68
MEMORY d~f
75b
61b
62b
51b
67
53b
75b
57b
58a
60b
MOTION DETECTOR
MOTION COMPENSATOR
PREDICTION MODE SELECTOR
TEXTURE ENCODER
VARIABLE-LENGTH ENCODER
94
85b
54b
59b
56b
44b
49b
45b
46b
52b
88b
SHAPE ENCODER
47b
46b
48b

Fig.12

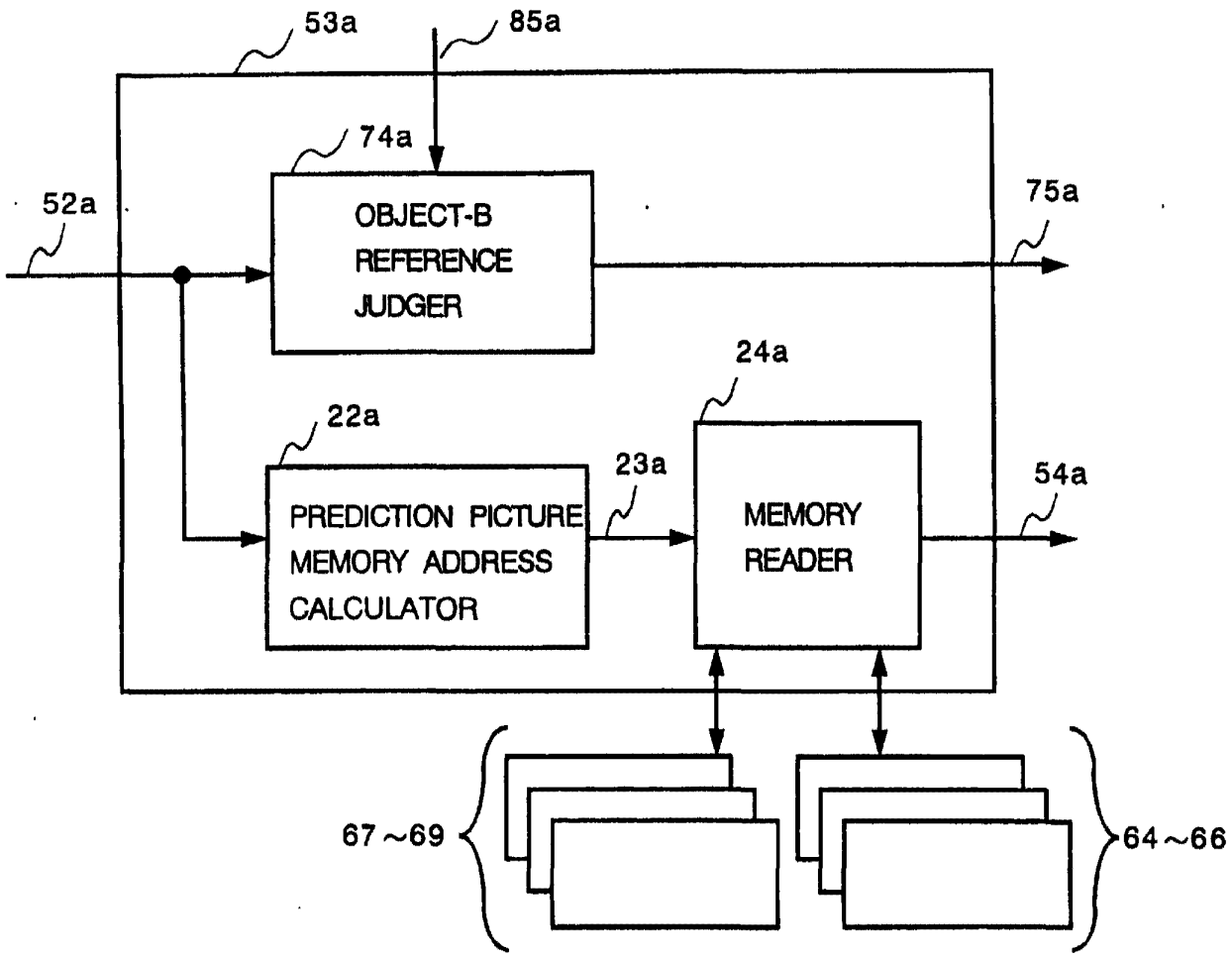

53a
85a
74a
52a
OBJECT-B
REFERENCE
JUDGER
75a
24a
22a
23a
54a
PREDICTION PICTURE
MEMORY ADDRESS
CALCULATOR
MEMORY
READER
67~69
64~66

Fig.13

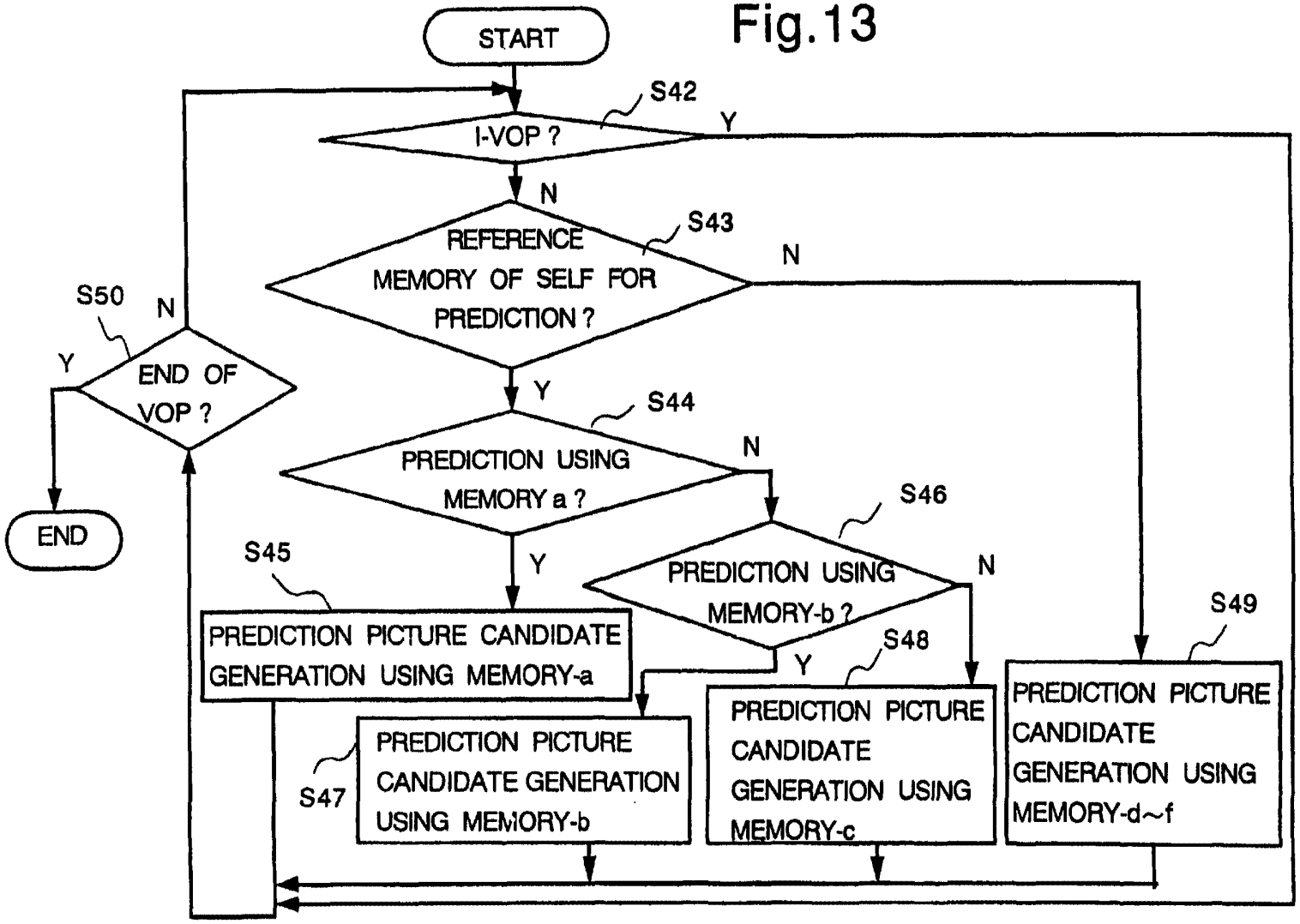

START
S42
I-VOP ?
Y
N
S43
REFERENCE MEMORY OF SELF FOR PREDICTION ?
N
Y
S44
PREDICTION USING MEMORY a ?
N
Y
S45
PREDICTION PICTURE CANDIDATE GENERATION USING MEMORY-a
S46
PREDICTION USING MEMORY-b ?
N
Y
S47
PREDICTION PICTURE CANDIDATE GENERATION USING MEMORY-b
S48
PREDICTION PICTURE CANDIDATE GENERATION USING MEMORY-c
S49
PREDICTION PICTURE CANDIDATE GENERATION USING MEMORY-d~f
S50
END OF VOP ?
N
Y
END

Fig.14

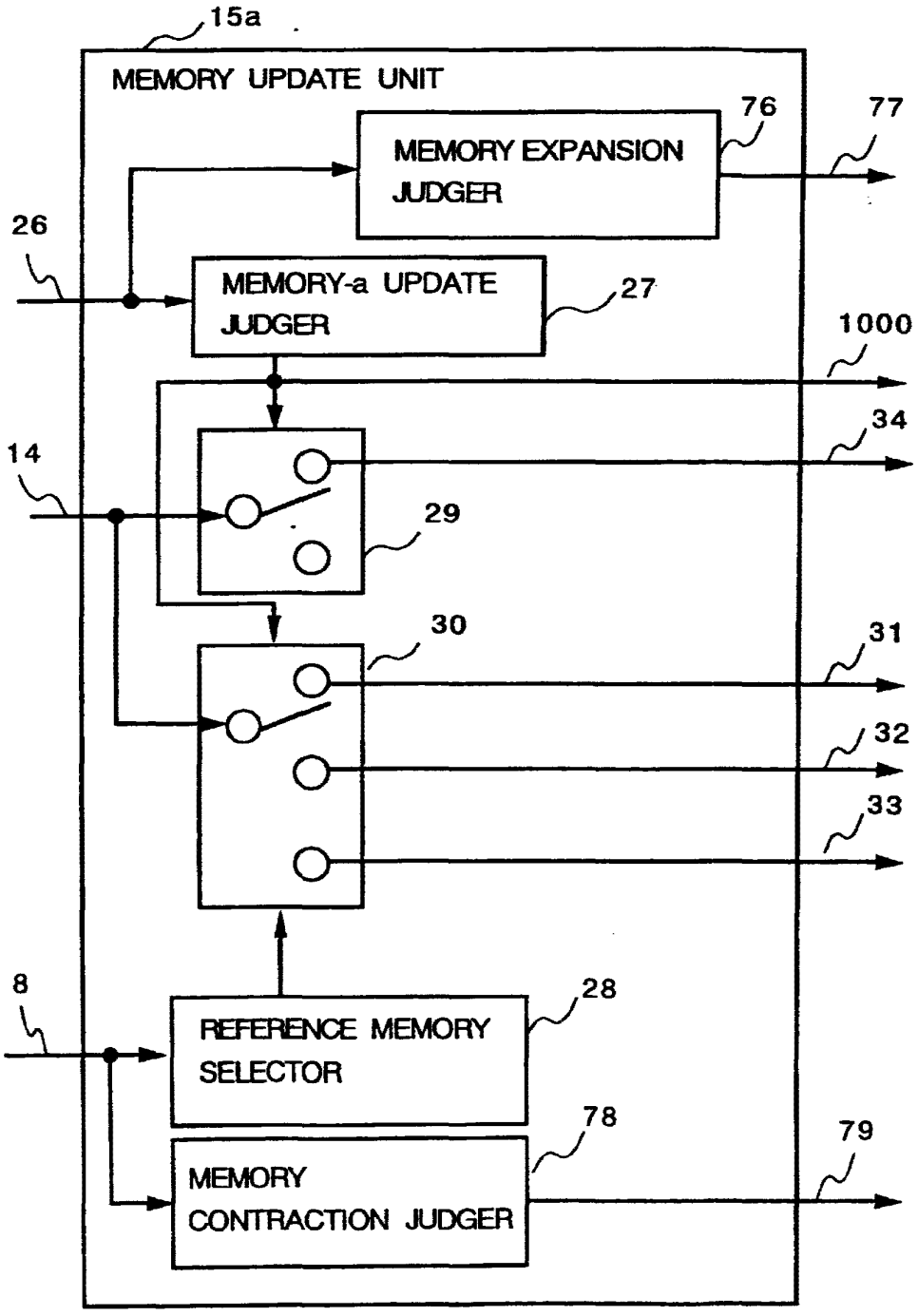

15a
MEMORY UPDATE UNIT
76
77
MEMORY EXPANSION JUDGER
26
MEMORY-a UPDATE JUDGER
27
1000
34
14
29
30
31
32
33
8
28
REFERENCE MEMORY SELECTOR
78
79
MEMORY CONTRACTION JUDGER

Fig.15

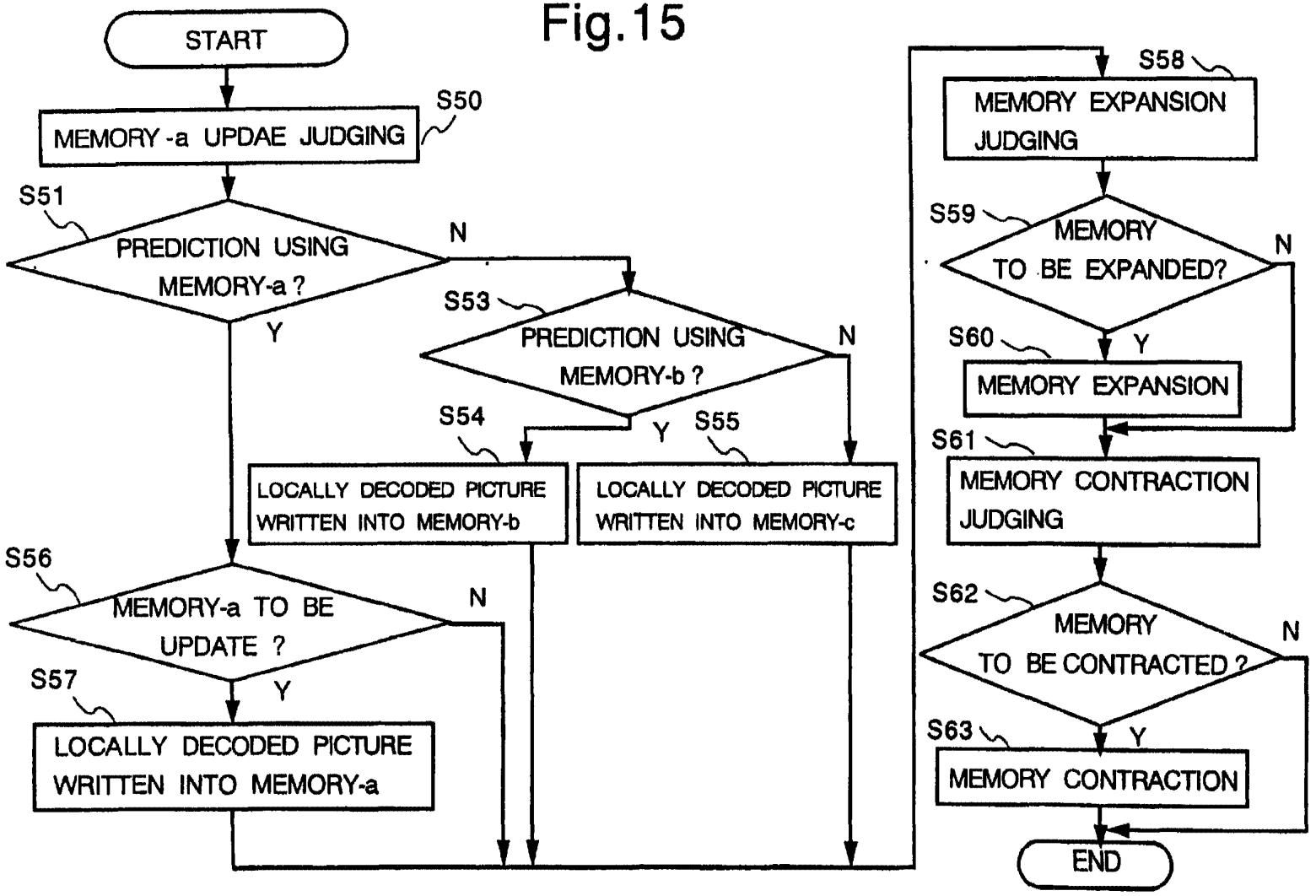

START
S50
MEMORY-a UPDAE JUDGING
S51
PREDICTION USING MEMORY-a ?
N
Y
S53
PREDICTION USING MEMORY-b ?
N
Y
S54
S55
LOCALLY DECODED PICTURE WRITTEN INTO MEMORY-b
LOCALLY DECODED PICTURE WRITTEN INTO MEMORY-c
S56
MEMORY-a TO BE UPDATE ?
N
Y
S57
LOCALLY DECODED PICTURE WRITTEN INTO MEMORY-a
S58
MEMORY EXPANSION JUDGING
S59
MEMORY TO BE EXPANDED?
N
Y
S60
MEMORY EXPANSION
S61
MEMORY CONTRACTION JUDGING
S62
MEMORY TO BE CONTRACTED ?
N
Y
S63
MEMORY CONTRACTION
END

Fig.16

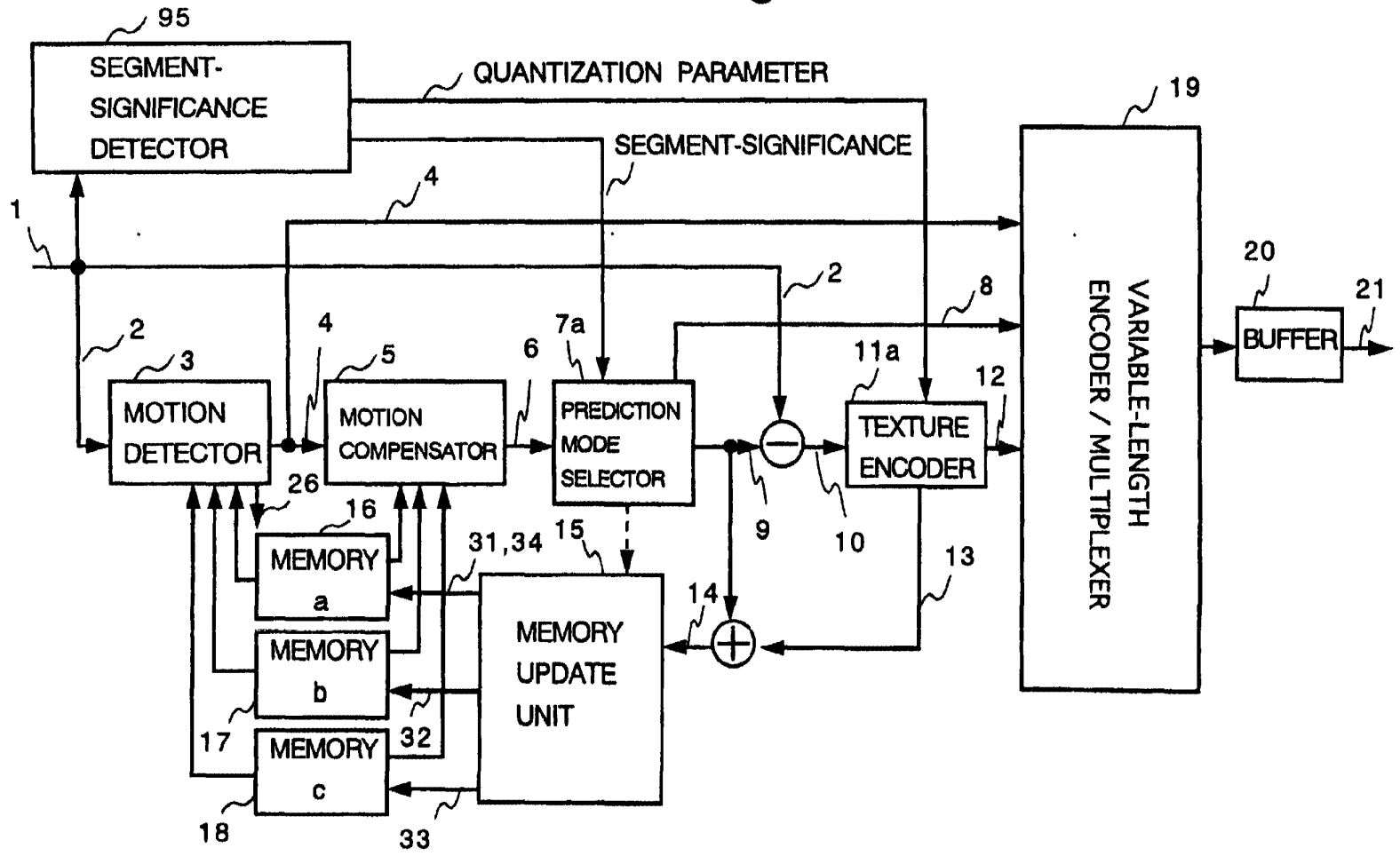

95
SEGMENT-SIGNIFICANCE DETECTOR
QUANTIZATION PARAMETER
SEGMENT-SIGNIFICANCE
19
VARIABLE-LENGTH ENCODER / MULTIPLEXER
20
BUFFER
21
1
4
2
8
2
3
MOTION DETECTOR
4
5
MOTION COMPENSATOR
6
7a
PREDICTION MODE SELECTOR
11a
TEXTURE ENCODER
12
26
16
MEMORY a
MEMORY b
MEMORY c
17
18
31,34
15
MEMORY UPDATE UNIT
32
33
9
10
13
14

Fig.17

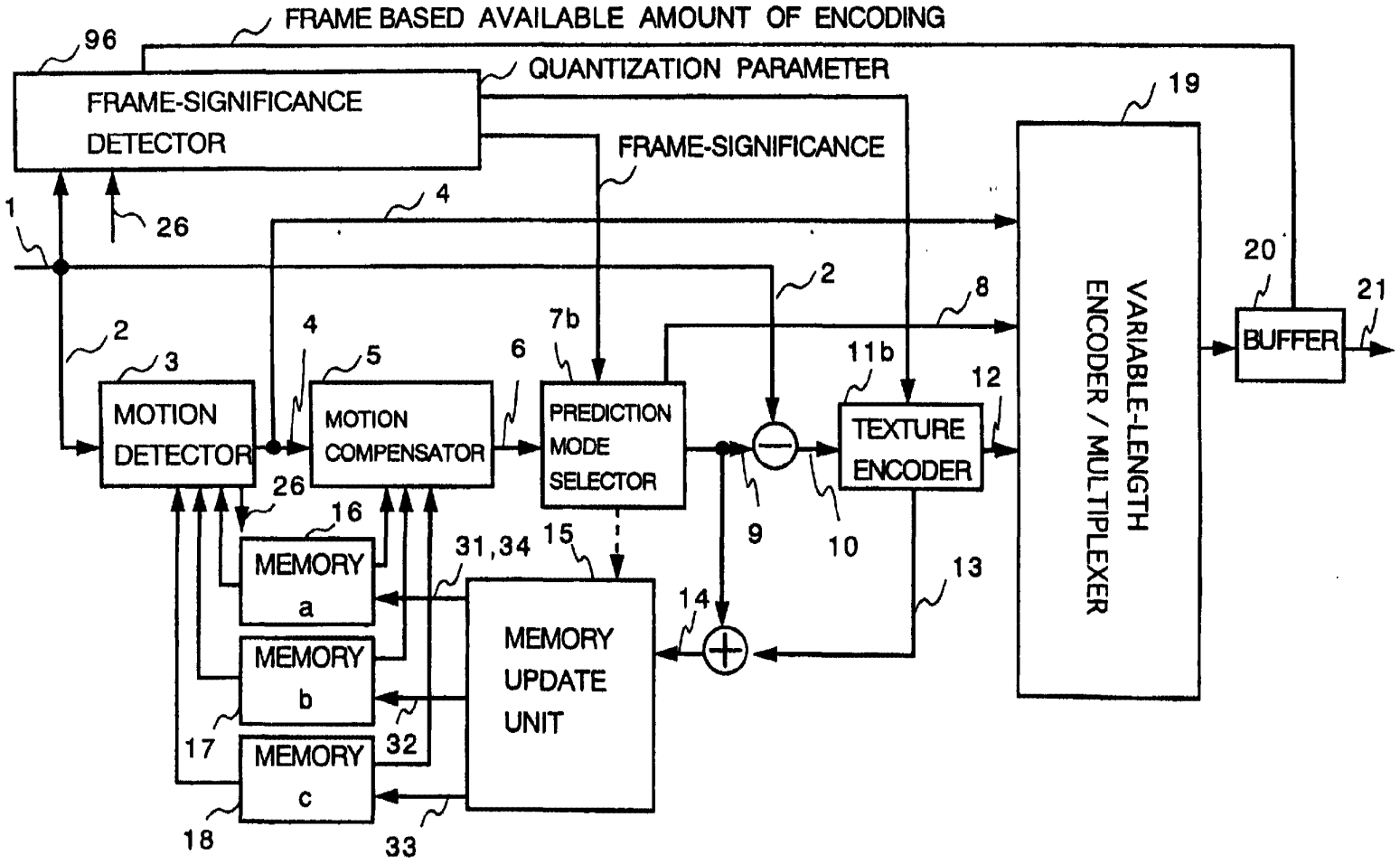

FRAME BASED AVAILABLE AMOUNT OF ENCODING
96
FRAME-SIGNIFICANCE DETECTOR
QUANTIZATION PARAMETER
FRAME-SIGNIFICANCE
19
1
26
4
2
8
20
21
BUFFER
VARIABLE-LENGTH ENCODER / MULTIPLEXER
2
3
MOTION DETECTOR
4
5
MOTION COMPENSATOR
6
7b
PREDICTION MODE SELECTOR
11b
12
TEXTURE ENCODER
26
16
MEMORY a
31,34
15
9
10
13
14
MEMORY b
MEMORY UPDATE UNIT
17
32
MEMORY c
18
33

Fig.18

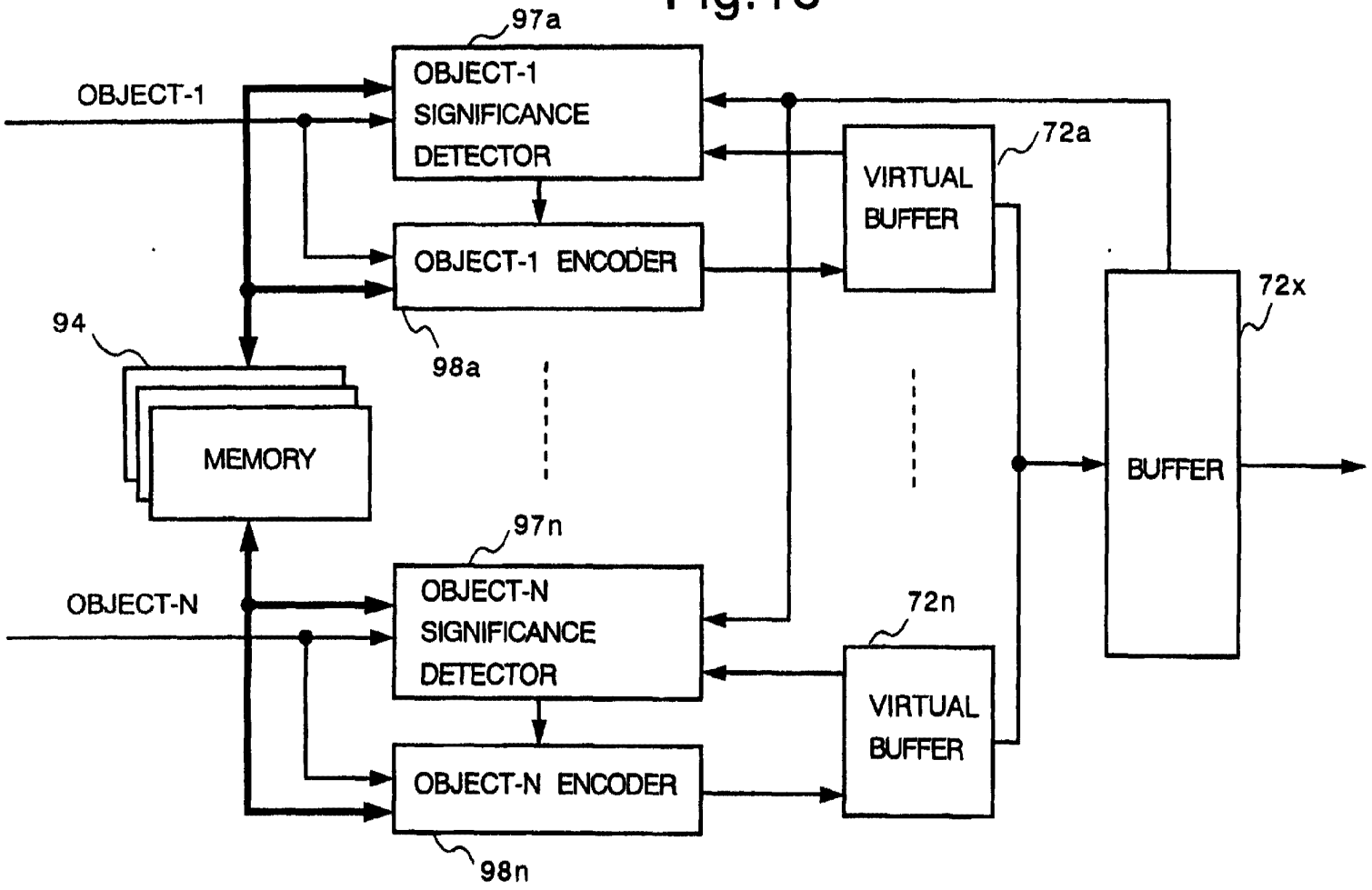

97a
OBJECT-1
OBJECT-1 SIGNIFICANCE DETECTOR
OBJECT-1 ENCODER
98a
72a
VIRTUAL BUFFER
72x
BUFFER
94
MEMORY
97n
OBJECT-N
OBJECT-N SIGNIFICANCE DETECTOR
OBJECT-N ENCODER
98n
72n
VIRTUAL BUFFER

Fig.19

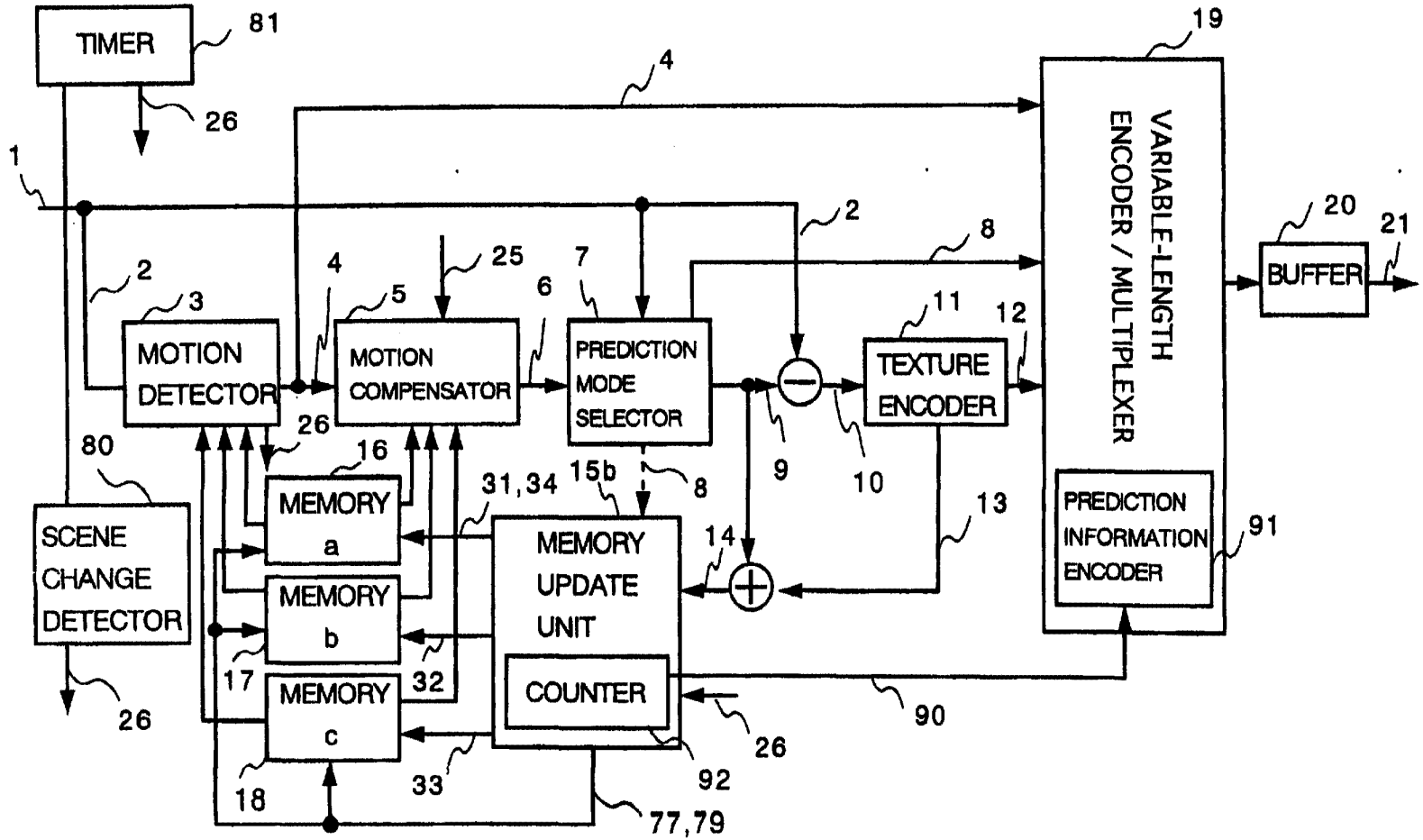
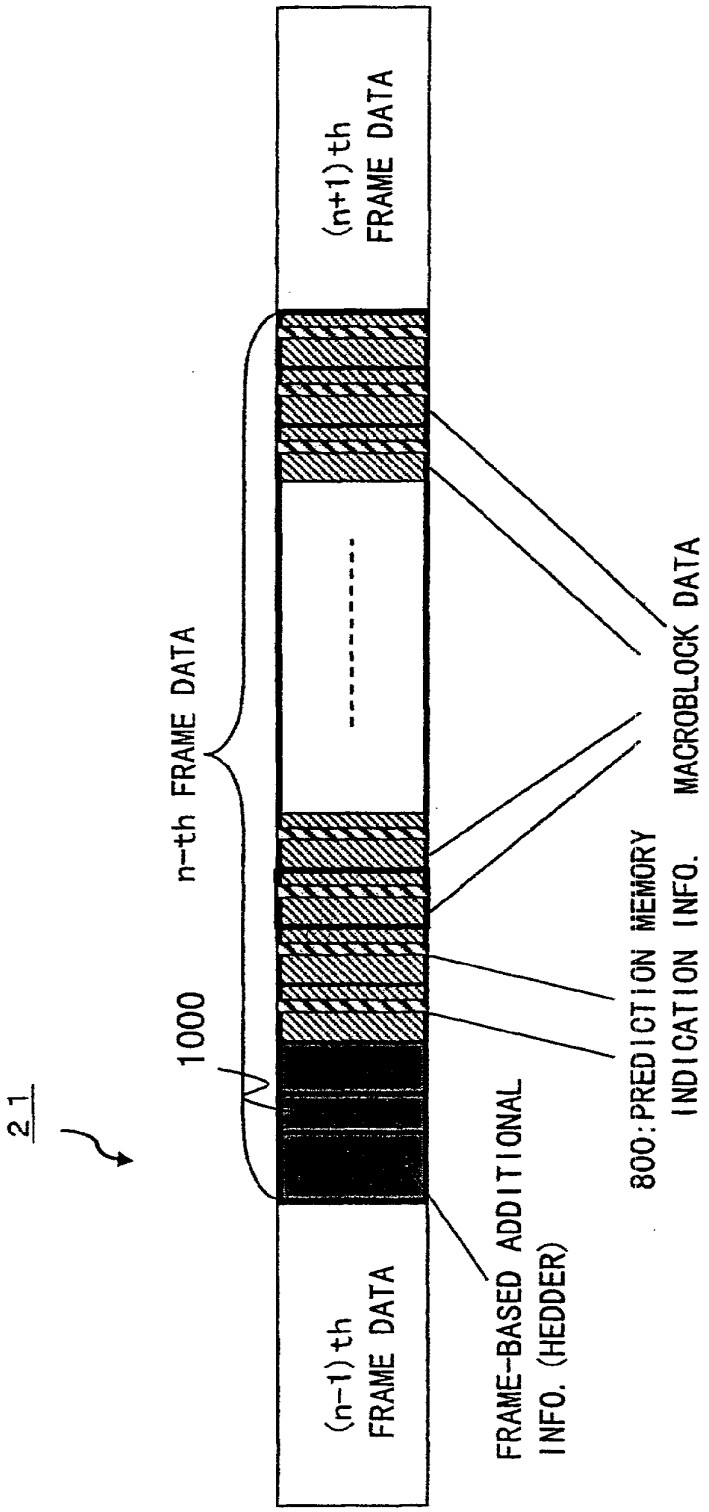

TIMER
81
26
4
19
VARIABLE-LENGTH ENCODER / MULTIPLEXER
1
2
8
20
21
BUFFER
2
3
4
5
25
6
7
11
12
MOTION DETECTOR
MOTION COMPENSATOR
PREDICTION MODE SELECTOR
TEXTURE ENCODER
26
80
16
31,34
15b
8
9
10
13
SCENE CHANGE DETECTOR
MEMORY a
MEMORY b
MEMORY c
MEMORY UPDATE UNIT
14
PREDICTION INFORMATION ENCODER
91
17
32
COUNTER
26
90
26
18
33
92
77,79

Fig. 20


21
1000
n-th FRAME DATA
(n+1)th FRAME DATA
(n-1)th FRAME DATA
FRAME-BASED ADDITIONAL INFO. (HEDDER)
800: PREDICTION MEMORY INDICATION INFO.
MACROBLOCK DATA

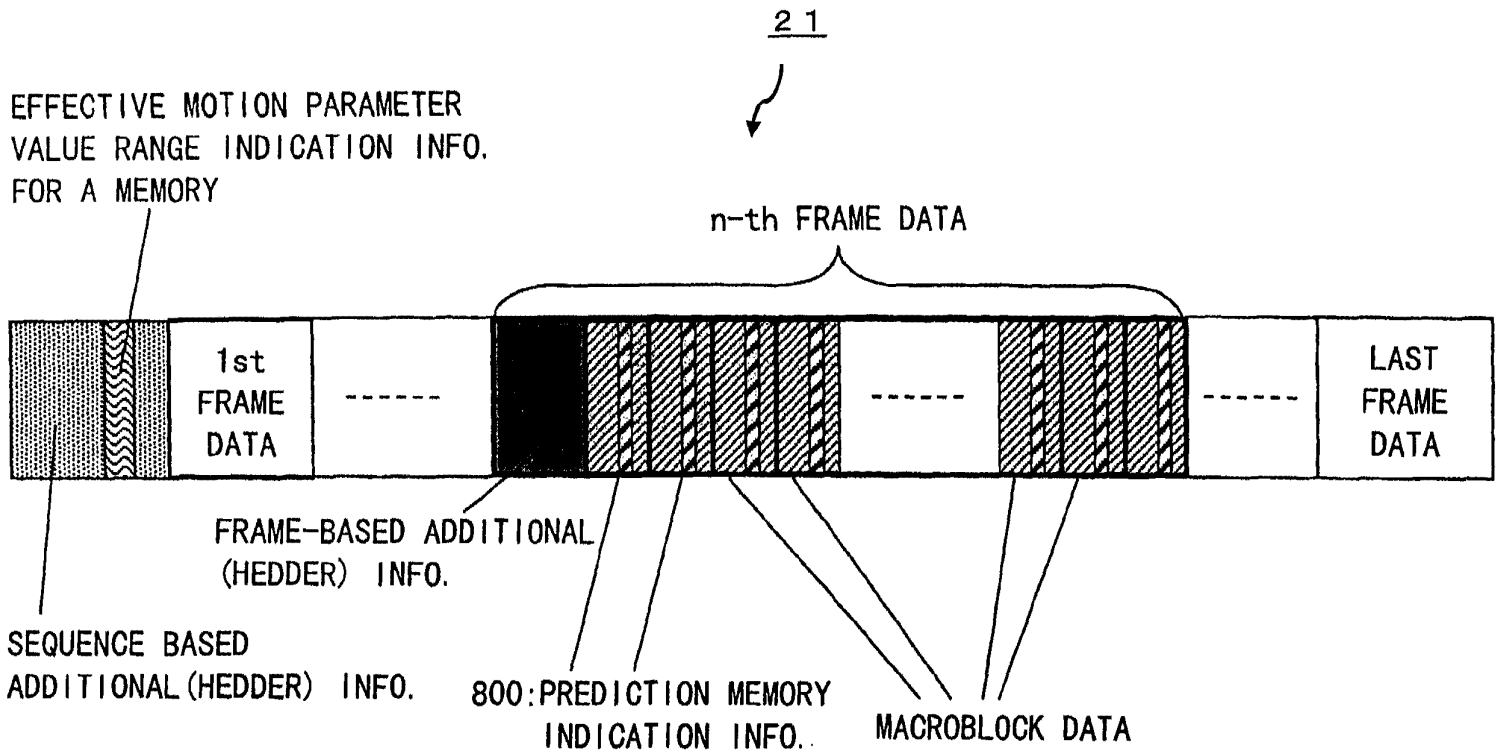

21
EFFECTIVE MOTION PARAMETER VALUE RANGE INDICATION INFO. FOR A MEMORY
n-th FRAME DATA
1st FRAME DATA
LAST FRAME DATA
FRAME-BASED ADDITIONAL (HEDDER) INFO.
SEQUENCE BASED ADDITIONAL (HEDDER) INFO.
800: PREDICTION MEMORY INDICATION INFO.
MACROBLOCK DATA

Fig. 21

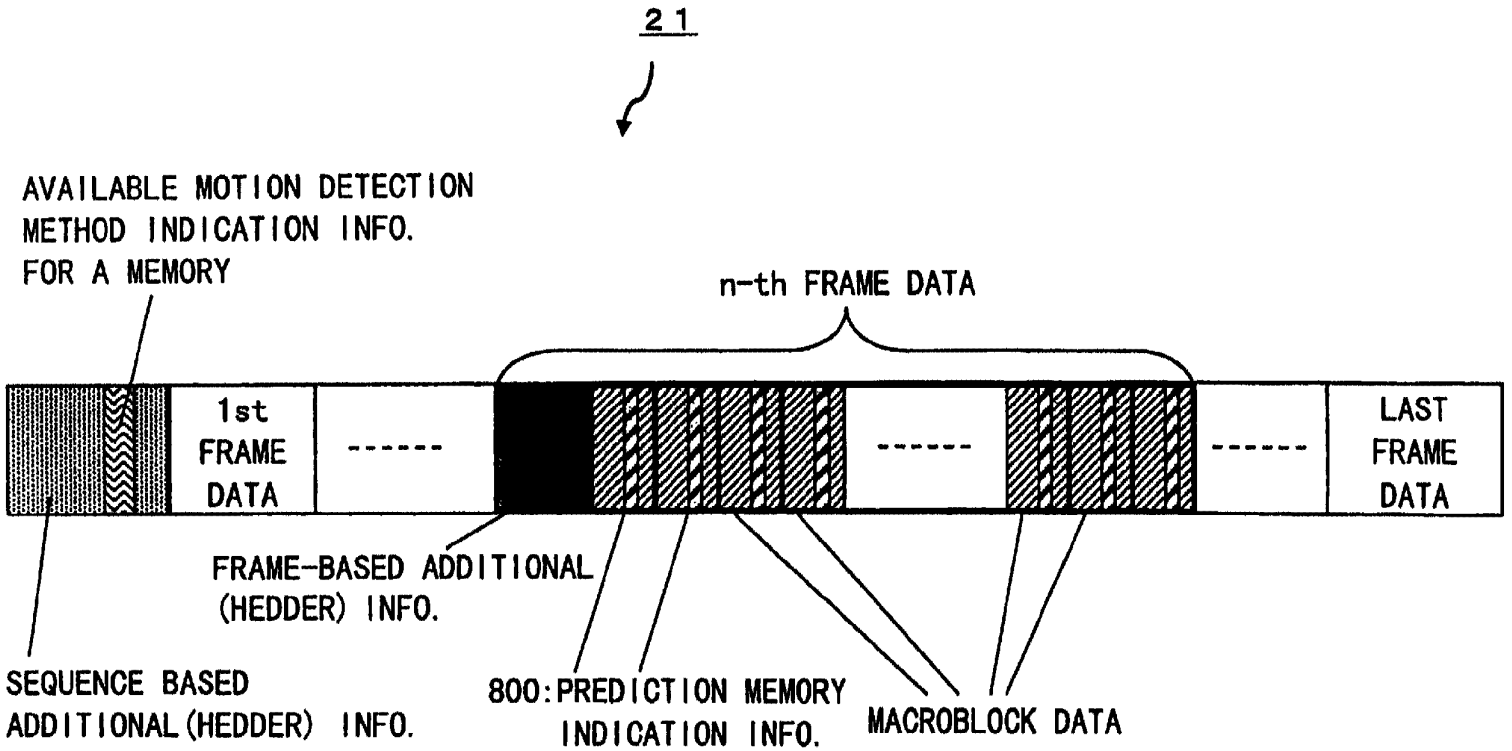
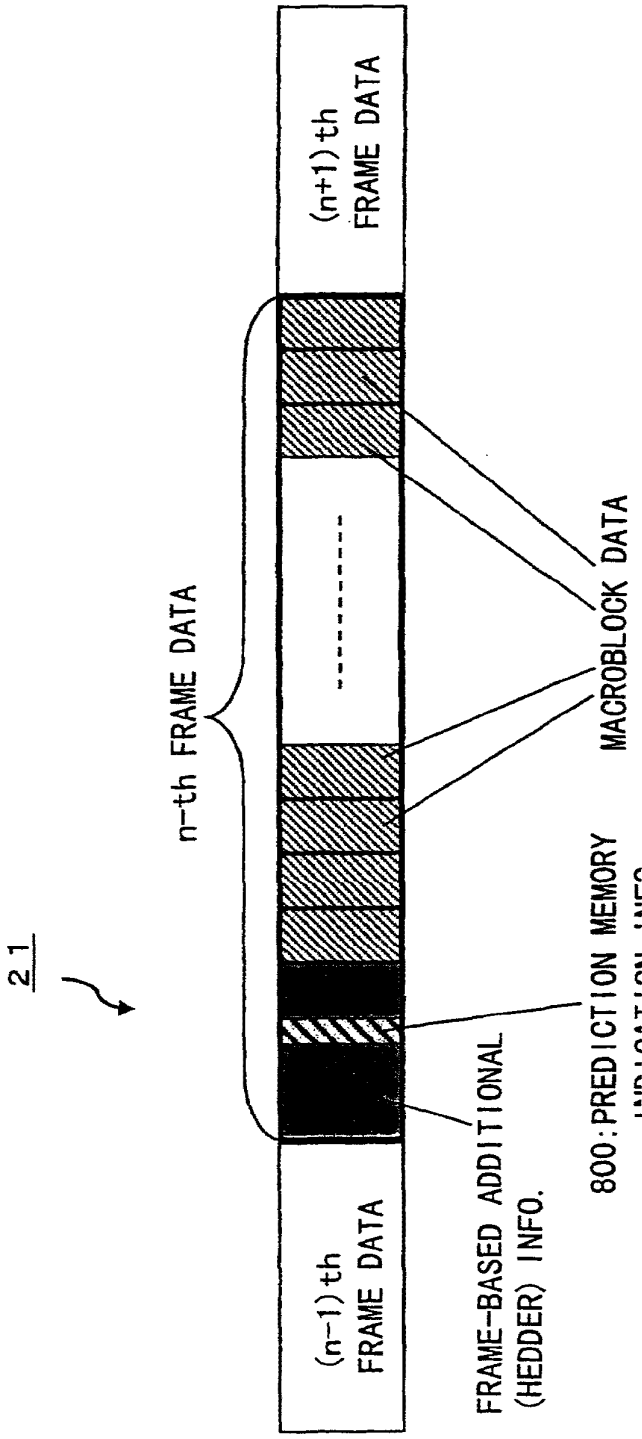

21
AVAILABLE MOTION DETECTION METHOD INDICATION INFO. FOR A MEMORY
n-th FRAME DATA
1st FRAME DATA
LAST FRAME DATA
FRAME-BASED ADDITIONAL (HEDDER) INFO.
SEQUENCE BASED ADDITIONAL (HEDDER) INFO.
800: PREDICTION MEMORY INDICATION INFO.
MACROBLOCK DATA

Fig. 22

Fig. 23


21
n-th FRAME DATA
(n+1)th FRAME DATA
(n-1)th FRAME DATA
MACROBLOCK DATA
FRAME-BASED ADDITIONAL (HEDDER) INFO.
800:PREDICTION MEMORY INDICATION INFO.

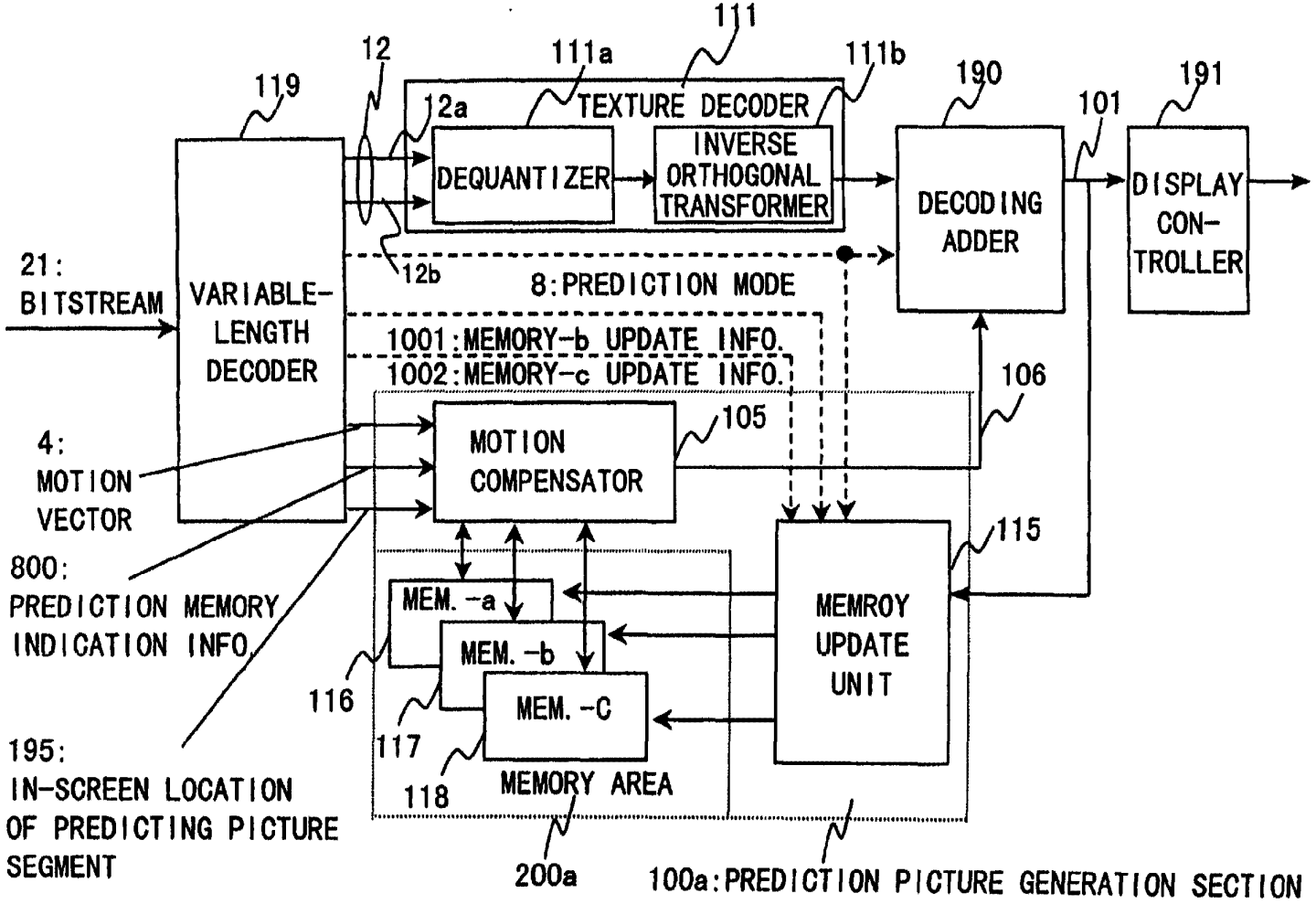


Fig. 24

EP 0 898 427 A1

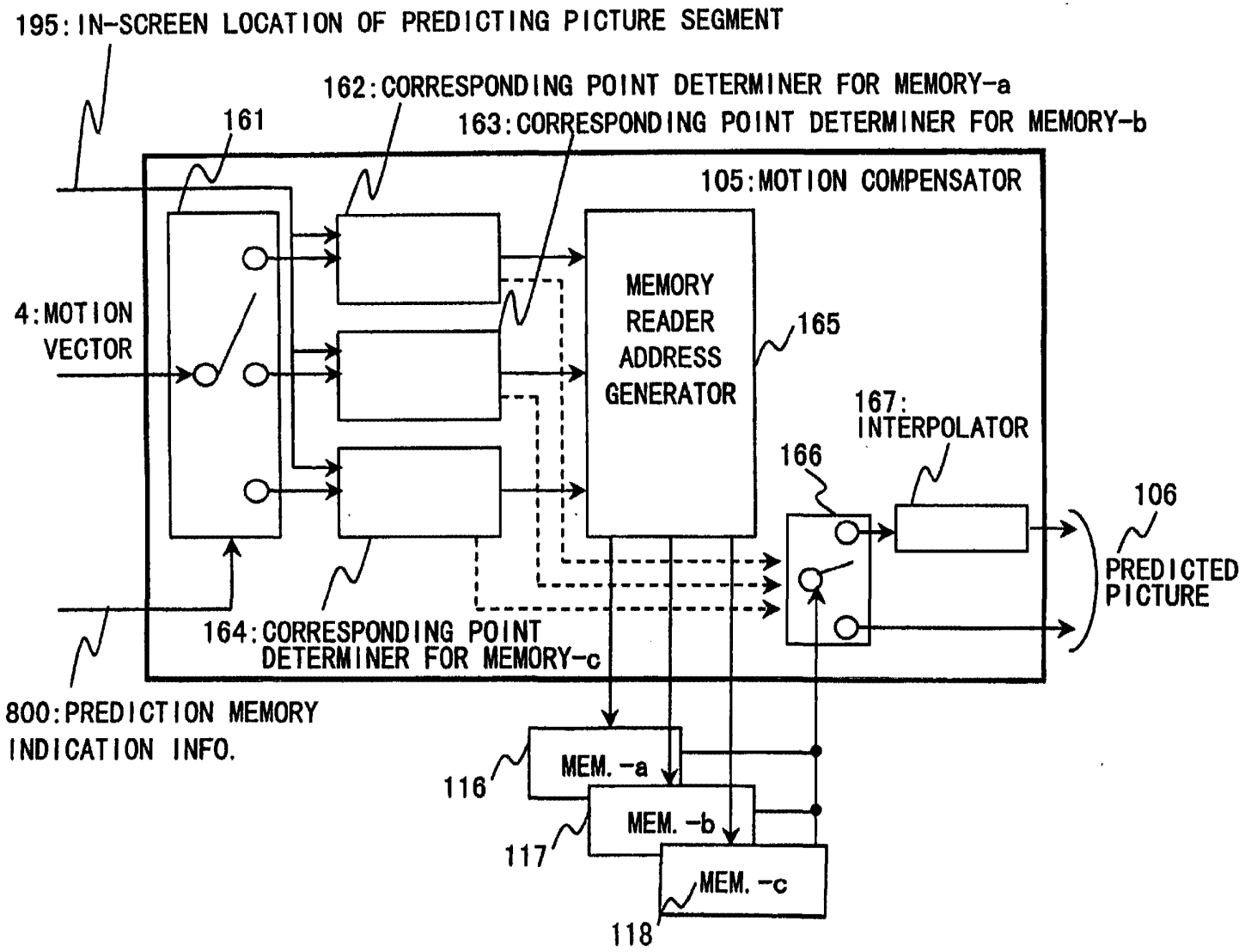


Fig. 25

EP 0 898 427 A1

Fig. 26

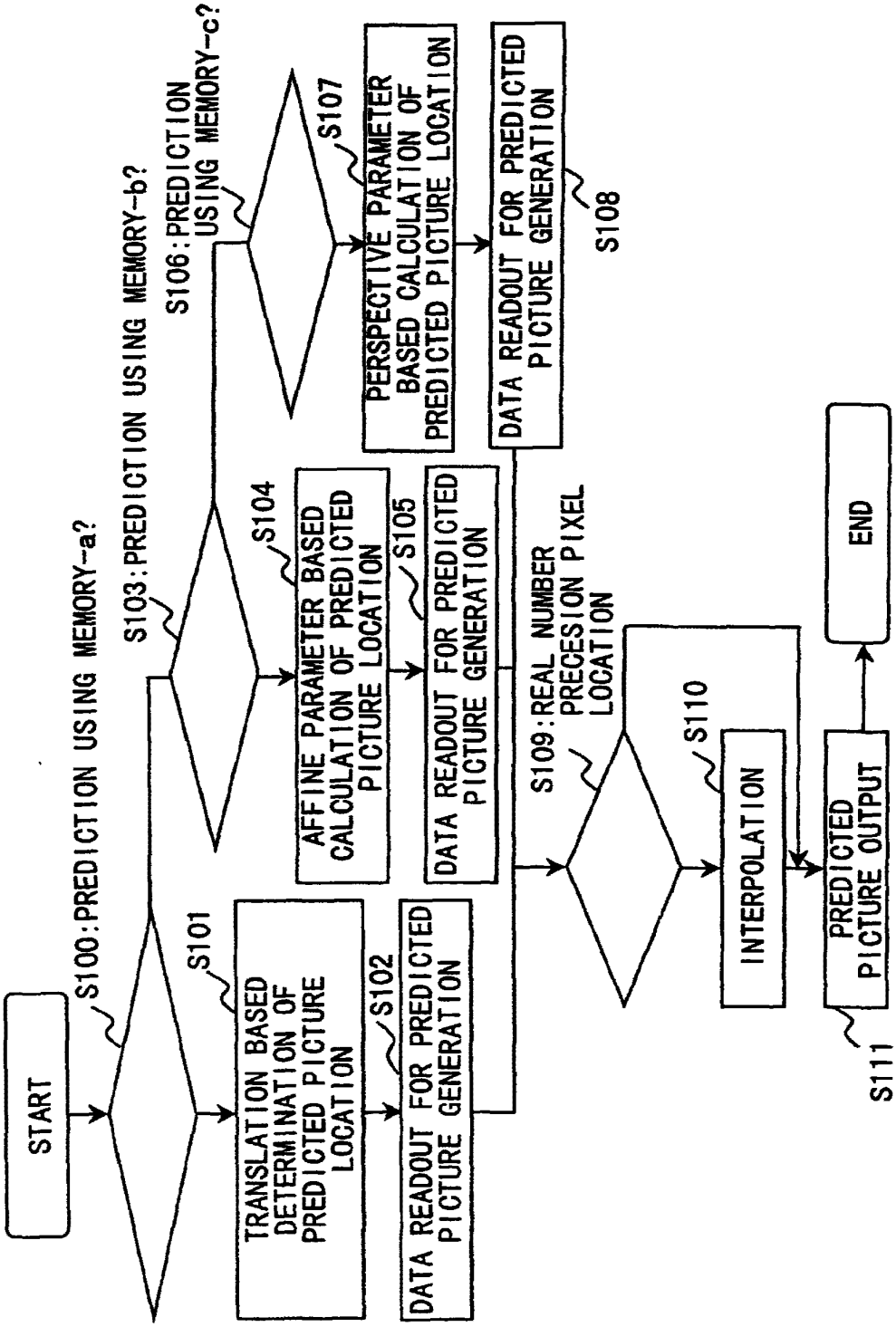

START
S100:PREDICTION USING MEMORY-a?
S101
TRANSLATION BASED DETERMINATION OF PREDICTED PICTURE LOCATION
S102
DATA READOUT FOR PREDICTED PICTURE GENERATION
S103:PREDICTION USING MEMORY-b?
S104
AFFINE PARAMETER BASED CALCULATION OF PREDICTED PICTURE LOCATION
S105
DATA READOUT FOR PREDICTED PICTURE GENERATION
S106:PREDICTION USING MEMORY-c?
S107
PERSPECTIVE PARAMETER BASED CALCULATION OF PREDICTED PICTURE LOCATION
DATA READOUT FOR PREDICTED PICTURE GENERATION
S108
S109:REAL NUMBER PRECESION PIXEL LOCATION
S110
INTERPOLATION
PREDICTED PICTURE OUTPUT
S111
END

Fig. 27

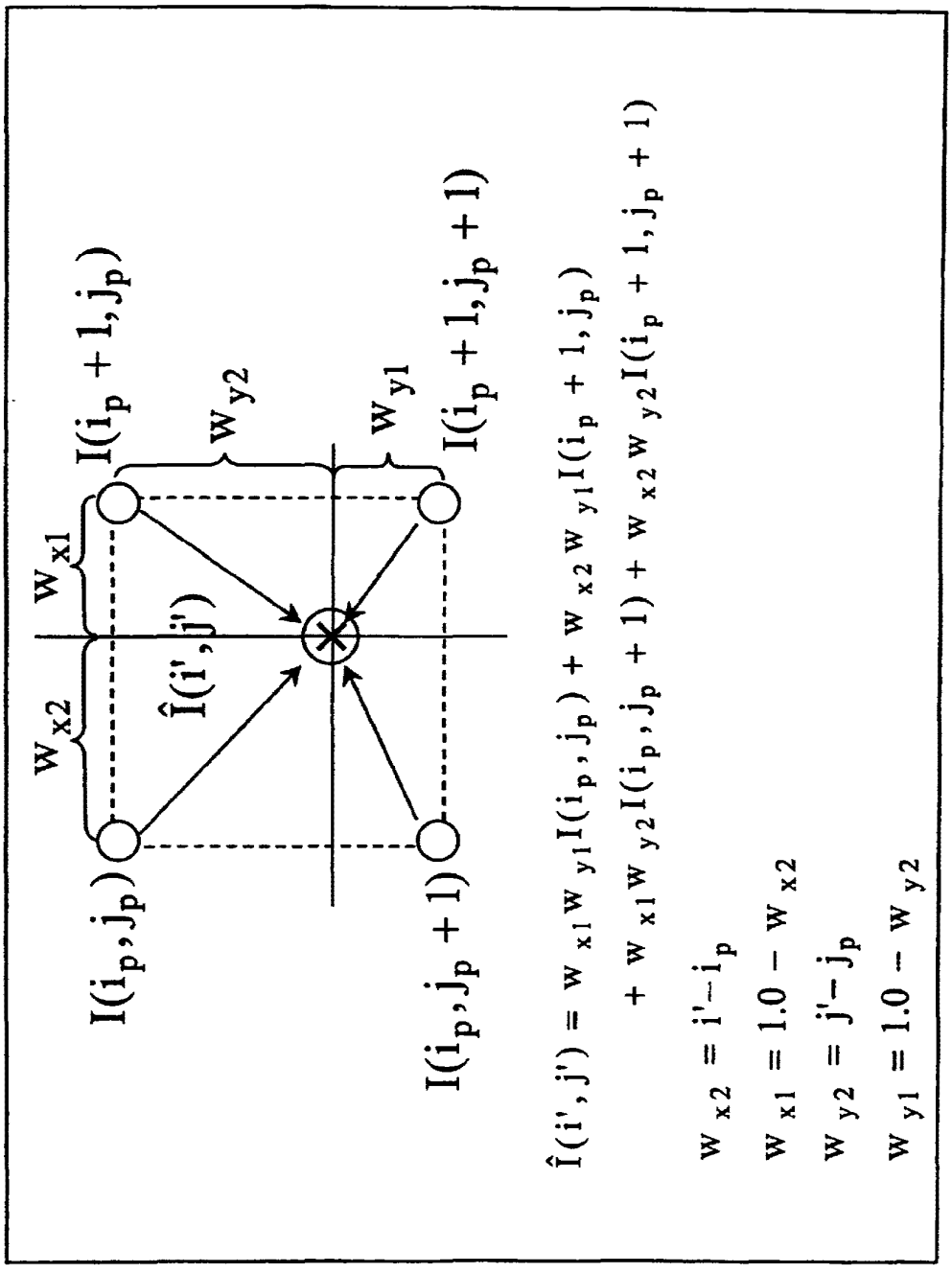

$I(i_p, j_p)$
$I(i_p + 1, j_p)$
$I(i_p, j_p + 1)$
$I(i_p + 1, j_p + 1)$
$\hat{I}(i', j')$
w_{x1}
w_{x2}
w_{y1}
w_{y2}
$\hat{I}(i', j') = w_{x1} w_{y1} I(i_p, j_p) + w_{x2} w_{y1} I(i_p + 1, j_p)$
$+ w_{x1} w_{y2} I(i_p, j_p + 1) + w_{x2} w_{y2} I(i_p + 1, j_p + 1)$
$w_{x2} = i' - i_p$
$w_{x1} = 1.0 - w_{x2}$
$w_{y2} = j' - j_p$
$w_{y1} = 1.0 - w_{y2}$

Fig. 28

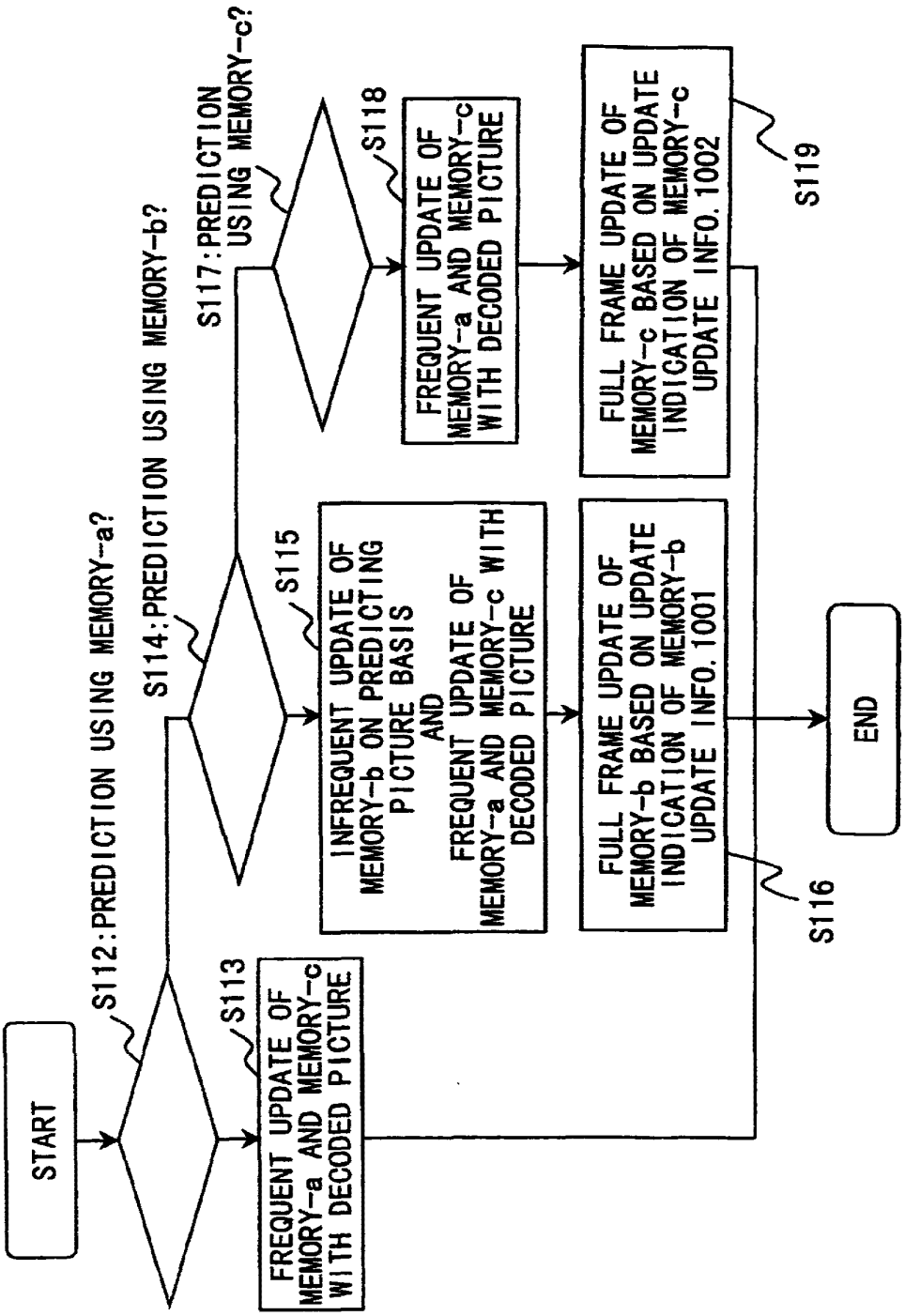

START
S112:PREDICTION USING MEMORY-a?
S113
FREQUENT UPDATE OF MEMORY-a AND MEMORY-c WITH DECODED PICTURE
S114:PREDICTION USING MEMORY-b?
S115
INFREQUENT UPDATE OF MEMORY-b ON PREDICTING PICTURE BASIS AND FREQUENT UPDATE OF MEMORY-a AND MEMORY-c WITH DECODED PICTURE
S116
FULL FRAME UPDATE OF MEMORY-b BASED ON UPDATE INDICATION OF MEMORY-b UPDATE INFO.1001
S117:PREDICTION USING MEMORY-c?
S118
FREQUENT UPDATE OF MEMORY-a AND MEMORY-c WITH DECODED PICTURE
S119
FULL FRAME UPDATE OF MEMORY-c BASED ON UPDATE INDICATION OF MEMORY-c UPDATE INFO.1002
END

Fig.29

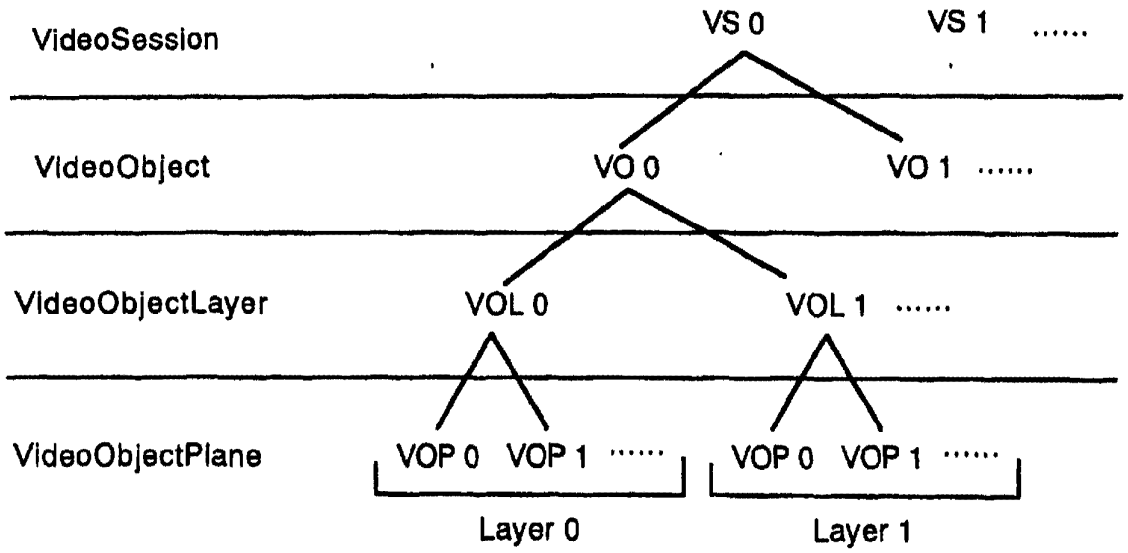

VideoSession
VS 0
VS 1 .....
VideoObject
VO 0
VO 1 .....
VideoObjectLayer
VOL 0
VOL 1 .....
VideoObjectPlane
VOP 0 VOP 1 .....
VOP 0 VOP 1 .....
Layer 0
Layer 1

Fig.30

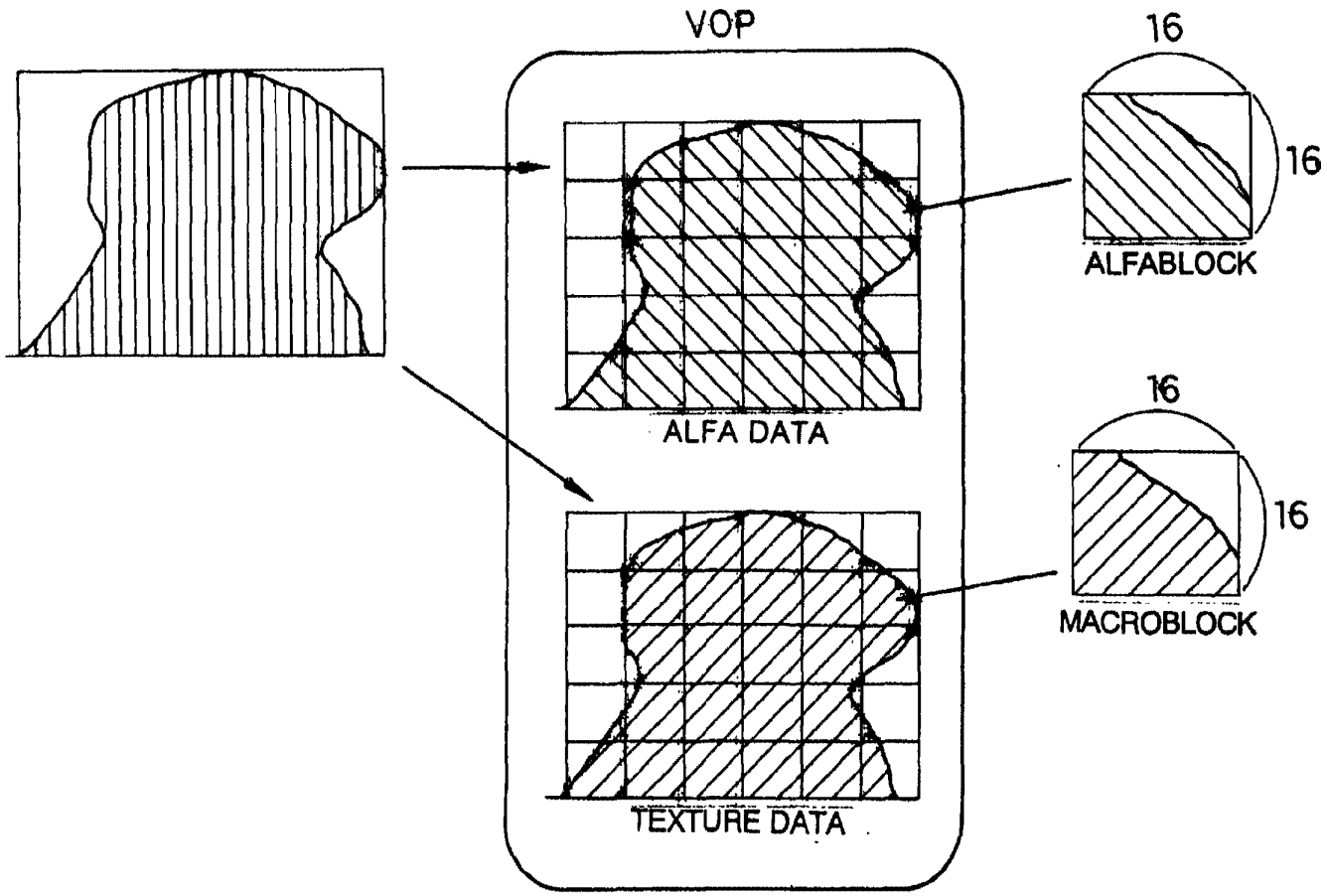


Fig.31

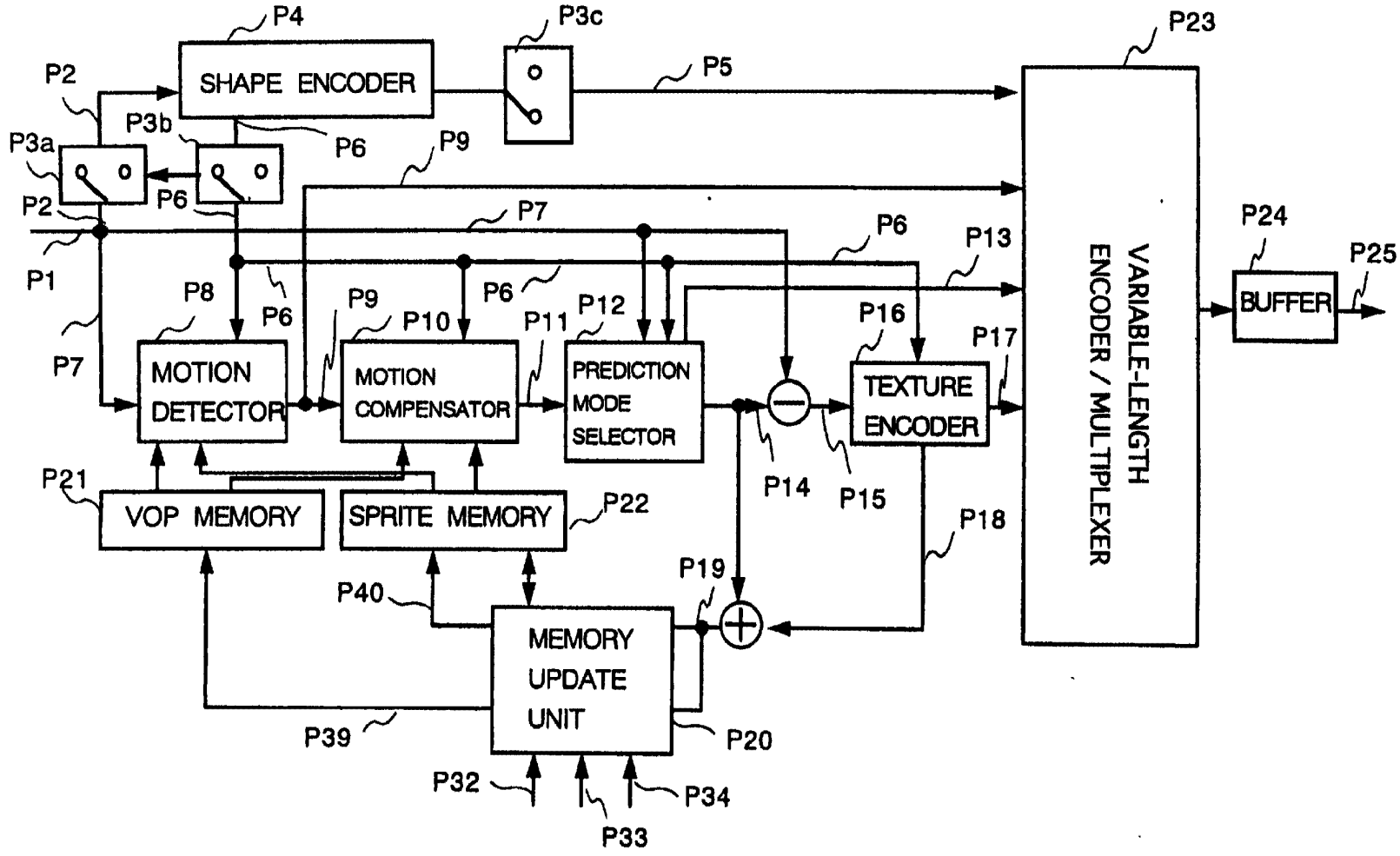

P4
SHAPE ENCODER
P3c
P23
P2
P5
P3b
P6
P9
P3a
P6
P2
P7
P6
P13
P24
P1
P6
P6
P25
P8
P9
P10
P11
P12
P16
P17
P7
MOTION DETECTOR
MOTION COMPENSATOR
PREDICTION MODE SELECTOR
TEXTURE ENCODER
VARIABLE-LENGTH ENCODER / MULTIPLEXER
BUFFER
P21
P14
P15
P22
VOP MEMORY
SPRITE MEMORY
P18
P19
P40
MEMORY UPDATE UNIT
P39
P20
P32
P34
P33

Fig.32

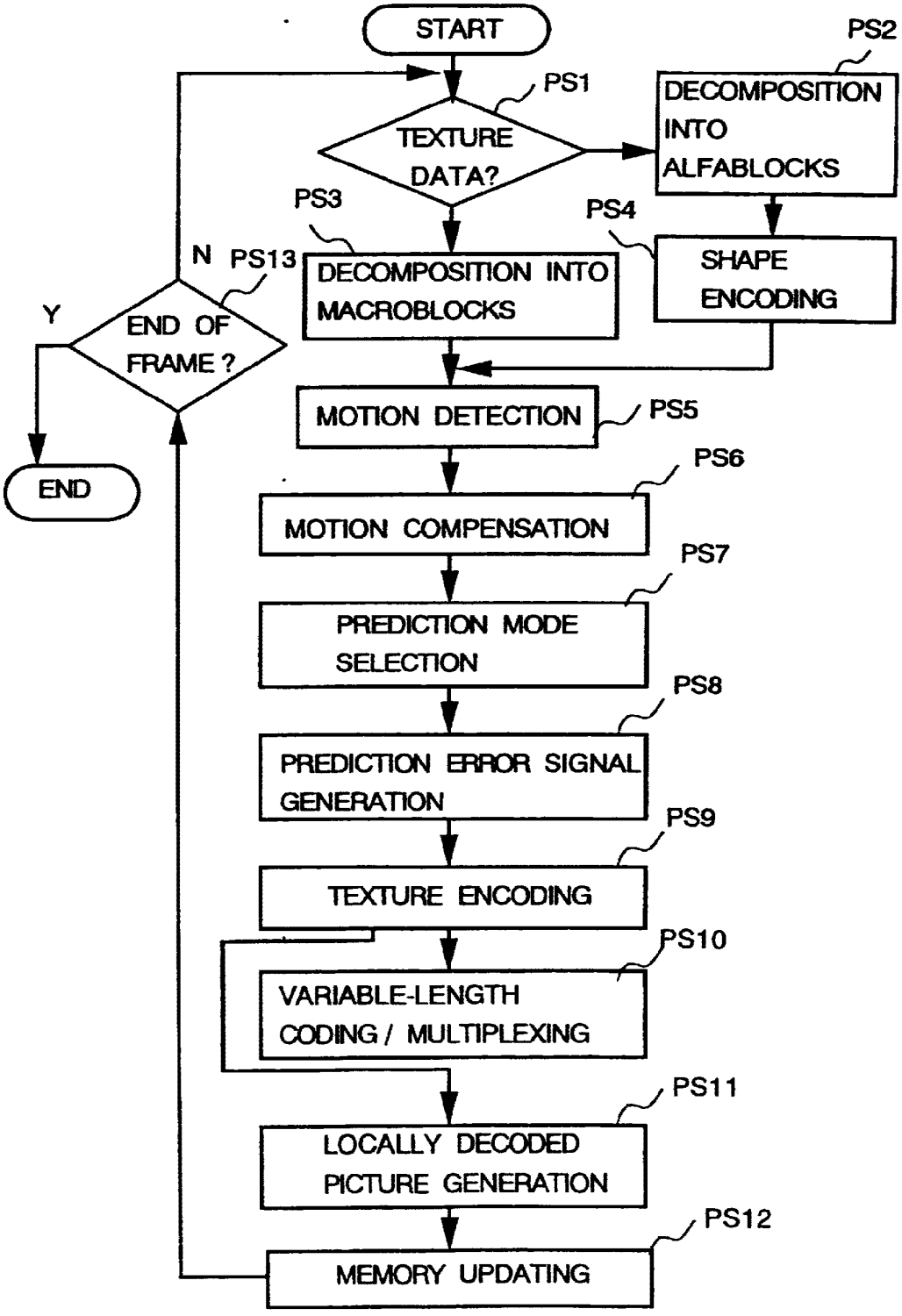

START
PS1
TEXTURE DATA?
PS2
DECOMPOSITION INTO ALFABLOCKS
PS3
DECOMPOSITION INTO MACROBLOCKS
PS4
SHAPE ENCODING
N
PS13
END OF FRAME ?
Y
END
MOTION DETECTION
PS5
MOTION COMPENSATION
PS6
PREDICTION MODE SELECTION
PS7
PREDICTION ERROR SIGNAL GENERATION
PS8
TEXTURE ENCODING
PS9
VARIABLE-LENGTH CODING / MULTIPLEXING
PS10
LOCALLY DECODED PICTURE GENERATION
PS11
MEMORY UPDATING
PS12

Fig.33

VOP TYPE	NO PREDICTION	PAST VOP BASED PREDICTION	PAST AND FUTURE VOP BASED PREDICTION	SPRITE BASED PREDICTION
I - VOP	○	×	×	×
P - VOP	○	○	×	×
B - VOP	○	○	○	×
SPRITE - VOP	○	○	×	○

Fig.34

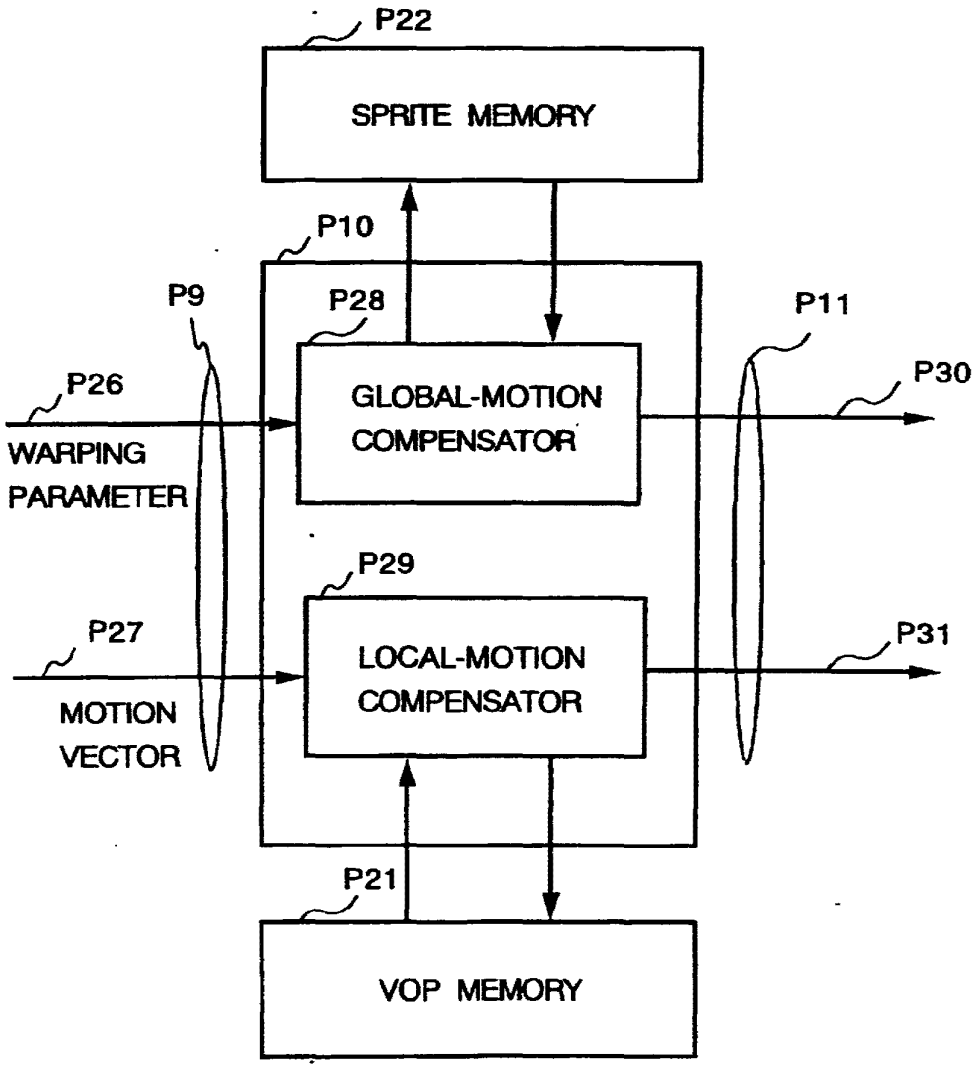

P22
SPRITE MEMORY
P10
P9
P28
P11
P26
GLOBAL-MOTION COMPENSATOR
P30
WARPING PARAMETER
P29
P27
LOCAL-MOTION COMPENSATOR
P31
MOTION VECTOR
P21
VOP MEMORY

Fig.35

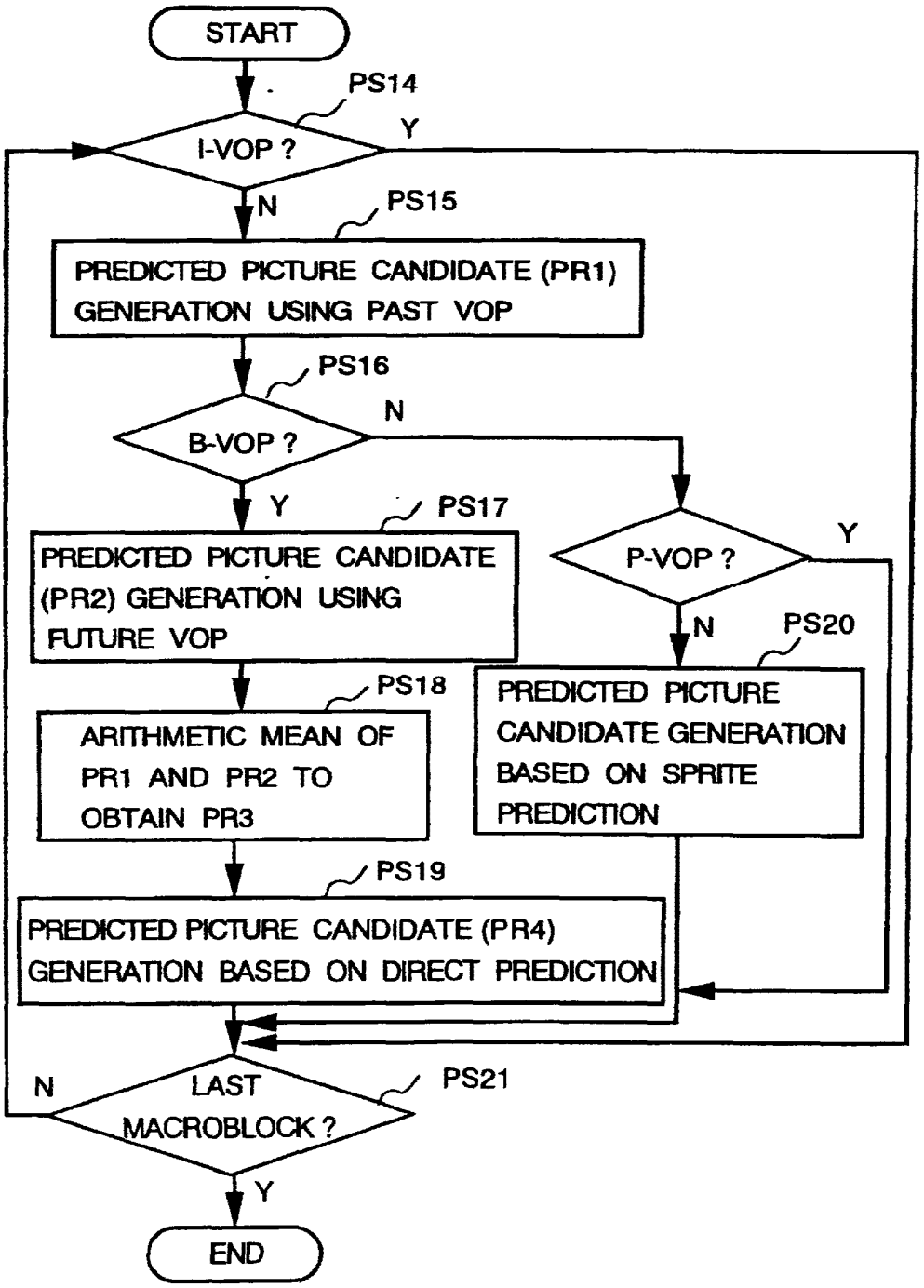

START
PS14
I-VOP ?
Y
N
PS15
PREDICTED PICTURE CANDIDATE (PR1) GENERATION USING PAST VOP
PS16
B-VOP ?
N
Y
PS17
PREDICTED PICTURE CANDIDATE (PR2) GENERATION USING FUTURE VOP
P-VOP ?
Y
N
PS20
PREDICTED PICTURE CANDIDATE GENERATION BASED ON SPRITE PREDICTION
PS18
ARITHMETIC MEAN OF PR1 AND PR2 TO OBTAIN PR3
PS19
PREDICTED PICTURE CANDIDATE (PR4) GENERATION BASED ON DIRECT PREDICTION
N
LAST MACROBLOCK ?
PS21
Y
END

Fig.36

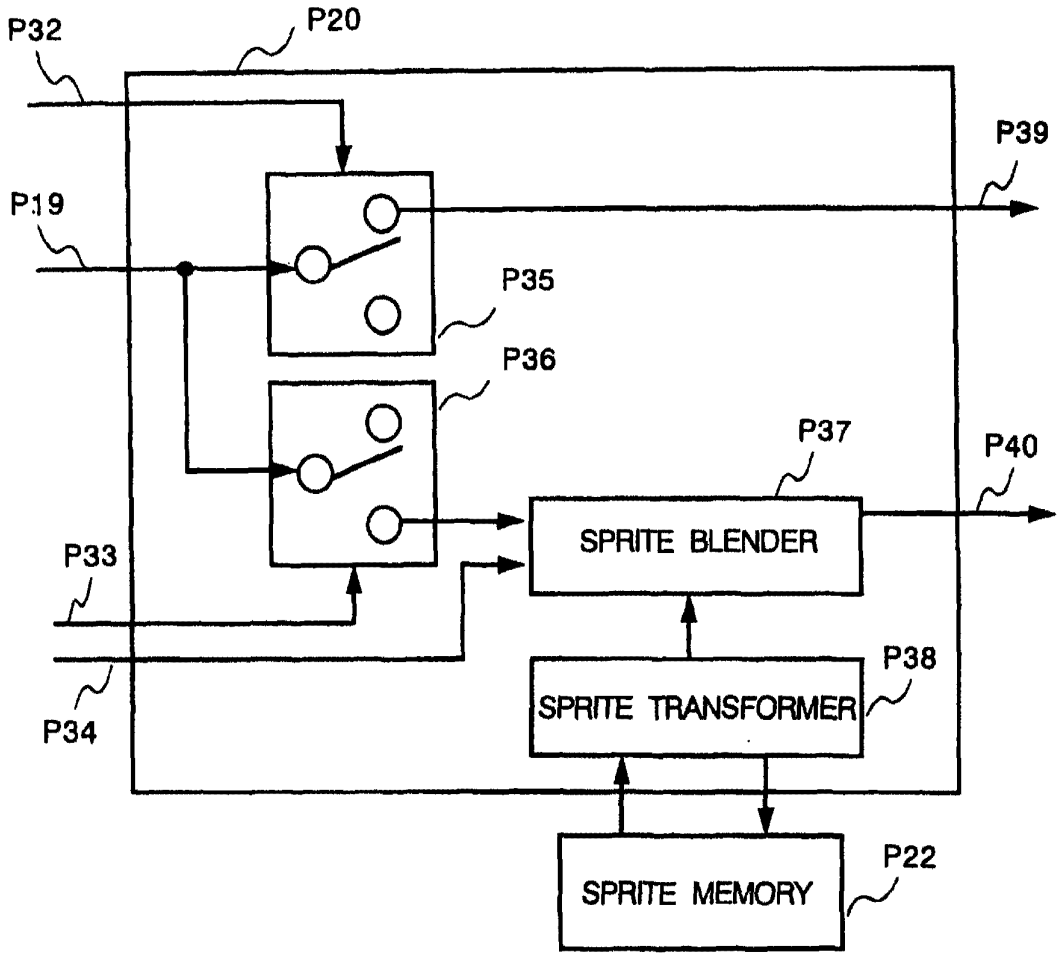
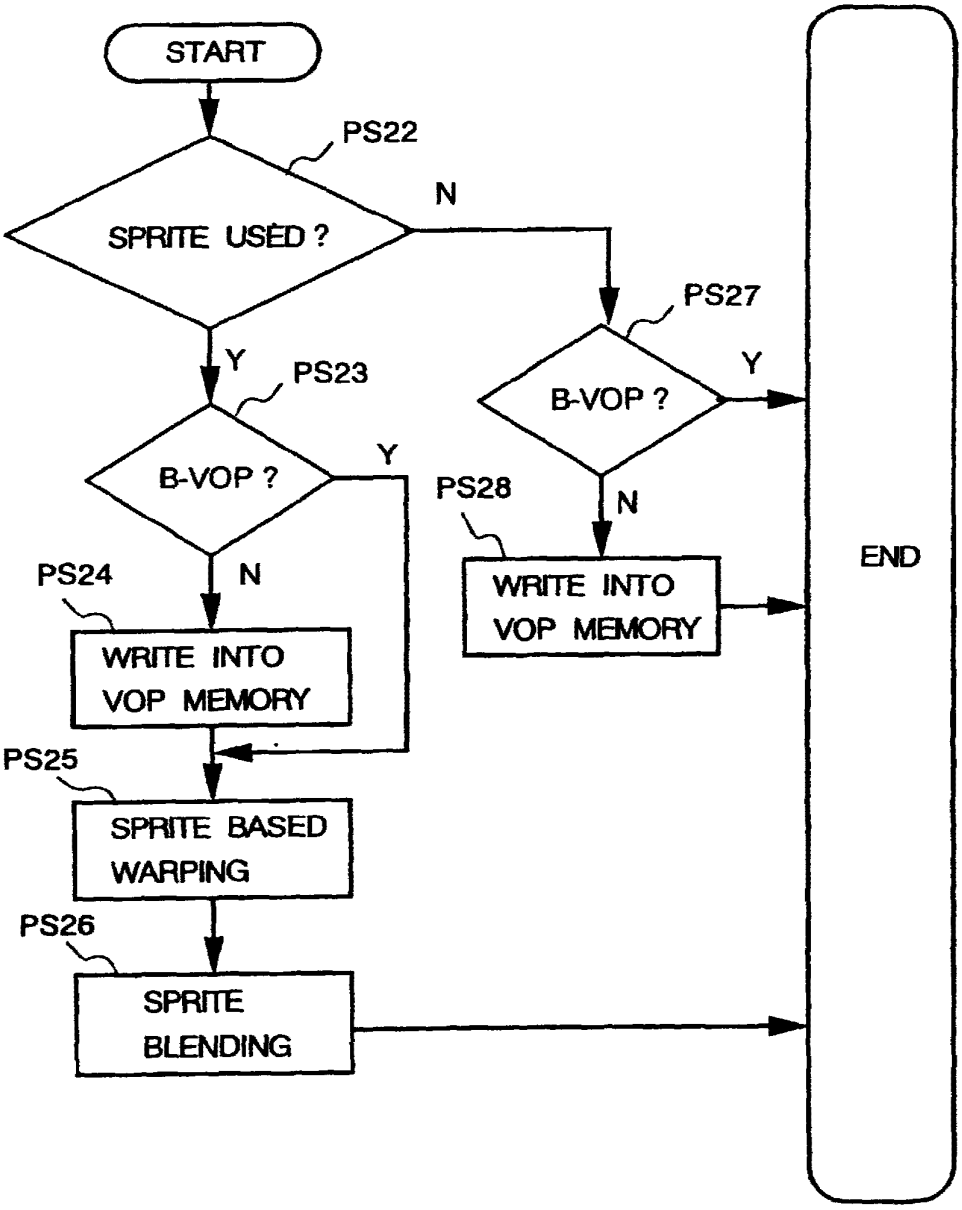

P32
P20
P39
P19
P35
P36
P37
P40
SPRITE BLENDER
P33
P38
SPRITE TRANSFORMER
P34
P22
SPRITE MEMORY

Fig.37


START
PS22
SPRITE USED ?
N
Y
PS23
B-VOP ?
Y
PS24
N
WRITE INTO VOP MEMORY
PS25
SPRITE BASED WARPING
PS26
SPRITE BLENDING
PS27
B-VOP ?
Y
PS28
N
WRITE INTO VOP MEMORY
END

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP98/00232

A. CLASSIFICATION OF SUBJECT MATTER
Int.Cl[6] H04N7/32

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
Int.Cl[6] H04N7/24-H04N7/68

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
Jitsuyo Shinan Koho 1957-1998
Kokai Jitsuyo Shinan Koho 1975-1998

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	JP, 8-172631, A (Matsushita Electric Industrial Co., Ltd.), July 2, 1996 (02. 07. 96) & EP, 719049, A2	1-3, 5, 9, 17-19, 21
Y		14-16
A		4, 6-8, 10-13, 20
Y	Hironao Nyozawa "Efficiency of Encoding in Conformity with MPEG-4 (in Japanese)", Preprints of 1996 Winter Meeting of Image Media Section, The Institute of Television Engineers of Japan, December 4, 1996 (04. 12. 96) (Tokyo) pages 39 to 44, particularly page 44, left column, lines 2 to 13	14-16
A		1-13, 17-21

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
April 21, 1998 (21. 04. 98)	May 12, 1998 (12. 05. 98)
Name and mailing address of the ISA/ Japanese Patent Office	Authorized officer
Facsimile No.	Telephone No.

Form PCT/ISA/210 (second sheet) (July 1992)

Electronic Acknowledgement Receipt

EFS ID:	7035879
Application Number:	11338934
International Application Number:	
Confirmation Number:	6610
Title of Invention:	Grouping of image frames in video coding
First Named Inventor/Applicant Name:	Miska Hannuksela
Customer Number:	76385
Filer:	Steven R. Funk
Filer Authorized By:	
Attorney Docket Number:	TPO.004.C2
Receipt Date:	18-FEB-2010
Filing Date:	25-JAN-2006
Time Stamp:	19:11:49
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Transmittal Letter	IDS-8-TPO004C2.pdf	54738 440dc377c432bfd5e70fff76ddca3c77468add40	no	2

Warnings:	
Information:	

2	Information Disclosure Statement (IDS) Filed (SB/08)	1449-8-TPO004C2.pdf	62014 7256dd08f9e1ff4b228f58983b044c089b45b2ad	no	2
Warnings:					
Information:					
This is not an USPTO supplied IDS fillable form					
3	NPL Documents	Elloumi.pdf	337636 7a38047d06aa5d3f7cd8e9e602921ea5c5b034ff	no	4
Warnings:					
Information:					
4	NPL Documents	Hannuksela.pdf	712317 50ee279cd1a1b1c7551108ba7cbed2abfb2c5507	no	14
Warnings:					
Information:					
5	NPL Documents	Soderquist.pdf	641172 08a2996ec505864cfaa3f7b8fa668d9f73f3a2ca	no	6
Warnings:					
Information:					
6	NPL Documents	EPSR-1.pdf	1448766 ac197f5e58085b037546485233f2c1327fd0f075	no	5
Warnings:					
Information:					
7	NPL Documents	EPSR-2.pdf	1732473 a0aa0d91756e7d74226ada4a85633e6f3969b1a9	no	5
Warnings:					
Information:					
8	Foreign Reference	EP0898427.pdf	2658400 a4a37ba6cebd0b966345da06aea44c900df3acab	no	60
Warnings:					
Information:					
9	NPL Documents	TAPA004PA-Allowance.pdf	507388 8687fa54e8edcd1b0f4307d369d6574904ccf80e	no	8
Warnings:					
Information:					
10	NPL Documents	TAPA004PA-OA09042009.pdf	328025 ee2eb11396b0d32275b729a41dcab7d02c43f190	no	7

Warnings:					
Information:					
11	NPL Documents	TAPA004PA-OA10082006.pdf	763601 b767b90add0d470accc93e5830f1b4af899fd807	no	16
Warnings:					
Information:					
12	NPL Documents	TAPA004PA-OA11092007.pdf	160508 fe1556ebcf37a1e60a1f35d6e46d29b7f2fe0fcd	no	4
Warnings:					
Information:					
13	NPL Documents	TAPA004PA-OA13032007.pdf	240311 b4f05eecd63d96dcc100f84be33ae6d606179a02	no	5
Warnings:					
Information:					
14	NPL Documents	TAPA004PA-OA16032006.pdf	169633 a6390f80af78ba4d9b79d8573a0af449252ae4bb	no	4
Warnings:					
Information:					
15	NPL Documents	TAPA004PA-OAR10082006.pdf	808657 c424432fc9383467ba6351f18e30eff8eac8e64d	no	16
Warnings:					
Information:					
16	NPL Documents	TAPA004PA-OAR-09042009.pdf	535372 2f1206b68db40e4ca8daad28ab010b904e168d6a	no	12
Warnings:					
Information:					
17	NPL Documents	TAPA004PA-OAR-13032007.pdf	543295 a8a2484b1702714bbea0af5d0d0a916eda91a5b7	no	12
Warnings:					
Information:					
18	NPL Documents	TAPA004PA-RR16032006.pdf	589717 dd9806cb4b2d7d65b32b33c965fa551ab50f97cb	no	13
Warnings:					
Information:					
			Total Files Size (in bytes):	12294023	

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111

If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371

If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office

If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant:	HANNUKSELA	Examiner:	Unassigned
Serial No.:	11/338,934	Group Art Unit:	2613
Filed:	January 25, 2006	Docket No.:	TPO.004.C2
Confirmation No.:	6610	Customer No.:	76385
Title:	GROUPING OF IMAGE FRAMES IN VIDEO CODING		

SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT (37 C.F.R. § 1.97(b))

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

With regard to the above-identified application, the items of information listed on the enclosed Form 1449 are brought to the attention of the Examiner.

This statement should be considered because it is submitted before the mailing date of a first Office Action on the merits for the above-identified application. Accordingly, no fee is due for consideration of the items listed on the enclosed Form 1449.

In accordance with 37 C.F.R. § 1.98(a)(2), only a copy of the foreign document or non-U.S. patent/application listed on the enclosed Form 1449 is provided. Also attached is a copy of the English translation of the Japanese Office Action.

Please note that any notations or markings on the attached documents do not reflect particular relevance, or lack thereof, to the present application, nor were they necessarily made by anyone affiliated with the prosecution of the present application.

No representation is made that a reference is "prior art" within the meaning of 35 U.S.C. §§ 102 and 103 and Applicant reserves the right, pursuant to 37 C.F.R. § 1.131 or otherwise, to

establish that the reference(s) are not "prior art." Moreover, Applicant does not represent that a reference has been thoroughly reviewed or that any relevance of any portion of a reference is intended.

Consideration of the items listed is respectfully requested. Pursuant to the provisions of MPEP § 609, it is requested that the Examiner return a copy of the attached Form 1449, marked as being considered and initialed by the Examiner, to the undersigned with the next official communication.

Respectfully submitted,

HOLLINGSWORTH & FUNK, LLC
8500 Normandale Lake Blvd., Suite 320
Minneapolis, MN 55437
952.854.2700

By: /Steven R. Funk/
Steven R. Funk
Reg. No. 37,830

FORM 1449* INFORMATION DISCLOSURE STATEMENT IN AN APPLICATION (Use several sheets if necessary)	Docket Number: TPO.004.C2	Application Number: 11/338,934
	Applicant: HANNUKSELA	
	Filing Date: 1/25/2006	Group Art Unit: 2613

U.S. PATENT DOCUMENTS

EXAMINER INITIAL	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	5122875	06/16/1992	Raychaudhuri et al.			
	5838265	11/17/1998	Adolph			
	5877812	03/02/1999	Krause et al.			
	6314139	11/06/2001	Koto et al.			

FOREIGN PATENT DOCUMENTS

	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	TRANSLATION	
						YES	NO

OTHER DOCUMENTS (Including Author, Title, Date, Pertinent Pages, Etc.)

	02/04/2010	Copy of Allowance Notification with translation dated 2/4/2010 from parallel Russian Application No. 2006110322, 17 pages.
	02/25/2010	Copy of translated Office Action dated 2/25/2010 from parallel Japanese Application No. 2006-120297, 5 pages.

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609; draw line through citation if not in conformance and not considered. Include copy of this form for next communication to the Applicant.

*Substitute Disclosure Statement Form (PTO-1449) Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Electronic Acknowledgement Receipt

EFS ID:	7316644
Application Number:	11338934
International Application Number:	
Confirmation Number:	6610
Title of Invention:	Grouping of image frames in video coding
First Named Inventor/Applicant Name:	Miska Hannuksela
Customer Number:	76385
Filer:	Steven R. Funk
Filer Authorized By:	
Attorney Docket Number:	TPO.004.C2
Receipt Date:	30-MAR-2010
Filing Date:	25-JAN-2006
Time Stamp:	15:45:16
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Transmittal Letter	IDS-9-TPO004C2.pdf	54555 9dff5f35f68114ec803240b8368ae5a9e9d88a1b	no	2

Warnings:	
Information:	

2	Information Disclosure Statement (IDS) Filed (SB/08)	1449-9-TPO004C2.pdf	52392 2b0412da4c73a2585f912f36d31dc7601983ba3e	no	1
Warnings:					
Information:					
This is not an USPTO supplied IDS fillable form					
3	NPL Documents	OA20100225.pdf	98113 56c0b5dbea6eaa02a0dd55397547b2bad6afd76f	no	5
Warnings:					
Information:					
4	NPL Documents	RUAllowance.pdf	2428420 81a2a0594f6e1200fcb4ddfe9abda3436c7a50f9	no	17
Warnings:					
Information:					
		Total Files Size (in bytes):	2633480		

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

<u>New Applications Under 35 U.S.C. 111</u>
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

<u>National Stage of an International Application under 35 U.S.C. 371</u>
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

<u>New International Application Filed with the USPTO as a Receiving Office</u>
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant:	HANNUKSELA	Examiner:	Unassigned
Serial No.:	11/338,934	Group Art Unit:	2613
Filed:	January 25, 2006	Docket No.:	TPO.004.C2
Confirmation No.:	6610	Customer No.:	76385
Title:	GROUPING OF IMAGE FRAMES IN VIDEO CODING		

SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT (37 C.F.R. § 1.97(b))

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

With regard to the above-identified application, the items of information listed on the enclosed Form 1449 are brought to the attention of the Examiner.

This statement should be considered because it is submitted before the mailing date of a first Office Action on the merits for the above-identified application. Accordingly, no fee is due for consideration of the items listed on the enclosed Form 1449.

In accordance with 37 C.F.R. § 1.98(a)(2), only a copy of the foreign document or non-U.S. patent/application listed on the enclosed Form 1449 is provided.

Please note that any notations or markings on the attached documents do not reflect particular relevance, or lack thereof, to the present application, nor were they necessarily made by anyone affiliated with the prosecution of the present application.

No representation is made that a reference is "prior art" within the meaning of 35 U.S.C. §§ 102 and 103 and Applicant reserves the right, pursuant to 37 C.F.R. § 1.131 or otherwise, to establish that the reference(s) are not "prior art." Moreover, Applicant does not represent that a

reference has been thoroughly reviewed or that any relevance of any portion of a reference is intended.

Consideration of the items listed is respectfully requested. Pursuant to the provisions of MPEP § 609, it is requested that the Examiner return a copy of the attached Form 1449, marked as being considered and initialed by the Examiner, to the undersigned with the next official communication.

Respectfully submitted,

HOLLINGSWORTH & FUNK, LLC
8500 Normandale Lake Blvd., Suite 320
Minneapolis, MN 55437
952.854.2700

By: /Steven R. Funk/
Steven R. Funk
Reg. No. 37,830

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
11/338,934	01/25/2006	Miska Hannuksela	TPO.004.C2	6610

76385 7590 05/25/2010

Hollingsworth & Funk
8500 Normandale Lake Blvd., Suite 320
Minneapolis, MN 55437

EXAMINER	
LEE, SIU M	

ART UNIT	PAPER NUMBER
2611	

MAIL DATE	DELIVERY MODE
05/25/2010	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary	Application No.	Applicant(s)	
	11/338,934	HANNUKSELA, MISKA	
	Examiner	Art Unit	
	SIU M. LEE	2611	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *25 January 2006*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-16* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-16* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *25 January 2006* is/are: a)☐ accepted or b)☒ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☒ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☒ Certified copies of the priority documents have been received in Application No. 10/306942.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date *See Continuation Sheet*.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

Continuation of Attachment(s) 3). Information Disclosure Statement(s) (PTO/SB/08), Paper No(s)/Mail Date :6/19/2006, 8/18/2006, 11/24/2006, 10/22/2007, 6/16/2008, 3/12/2009, 4/20/2009, 2/08/2010, 2/18/2010, 3/30/2010.

DETAILED ACTION

Claim Objections

1. Claims 1-4, 5-8, and 10-12 are objected to because of the following informalities:

Claim 1, line 4 recites “the previous reference image frame”; there is a lack of antecedent basis for this limitation.

Claims 5, line 4 recites “the previous reference image frame”; there is a lack of antecedent basis for this limitation.

Claim 10, line 3 recites “the previous reference image frame”; there is a lack of antecedent basis for this limitation.

Appropriate correction is required.

Drawings

2. The drawings are objected to because Figure 1 does not have sufficiently descriptive labels. Blank boxes in drawings should be labeled descriptively unless it is a well-known component. For example, the blank box, "100" should be labeled as "Multimedia Source". Corrected drawing sheets in compliance with 37 CFR 1.121(d) are required in reply to the Office action to avoid abandonment of the application. Any amended replacement drawing sheet should include all of the figures appearing on the immediate prior version of the sheet, even if only one figure is being amended. The figure or figure number of an amended drawing should not be labeled as “amended.” If a drawing figure is to be canceled, the appropriate figure must be removed from the

replacement sheet, and where necessary, the remaining figures must be renumbered and appropriate changes made to the brief description of the several views of the drawings for consistency. Additional replacement sheets may be necessary to show the renumbering of the remaining figures. Each drawing sheet submitted after the filing date of an application must be labeled in the top margin as either "Replacement Sheet" or "New Sheet" pursuant to 37 CFR 1.121(d). If the changes are not accepted by the examiner, the applicant will be notified and informed of any required corrective action in the next Office action. The objection to the drawings will not be held in abeyance.

Claim Rejections - 35 USC § 101

3. 35 U.S.C. 101 reads as follows:

> Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

Claims 5-8 are rejected under 35 U.S.C. 101 because the claimed invention is directed to non-statutory subject matter. The claim is NOT to a process, machine, manufacture or composition of matter. The claimed "video encoder" is non-structural per se and because the specification discloses a "software" embodiment (page 5, paragraph 0014 and page 33, paragraph 0097). Therefore, a reasonable interpretation in light of the specification leads to the conclusion that the claim video encoder as pure software (further confirm in claim 9 that computer program code for encoding a compressed video sequence, claim 9, line 7), which does not fall within the definition of a process, machine, manufacture or composition of matter.

Claim 13 is rejected under 35 U.S.C. 101 because the claimed invention is directed to non-statutory subject matter. The claim is NOT to a process, machine, manufacture or composition of matter. The claimed "video decoder" is non-structural per se and because the specification discloses a "software" embodiment (page 5, paragraph 0014 and page 33, paragraph 0097). Therefore, a broadest reasonable interpretation in light of the specification leads to the conclusion that the claim video decoder as pure software (further confirm in claim 9 that computer program code for decoding from the video sequence, claim 14, line 7), which does not fall within the definition of a process, machine, manufacture or composition of matter.

Claim 9 is rejected under 35 U.S.C. 101 because the claimed invention is directed to non-statutory subject matter. The claim is NOT to a process, machine, manufacture or composition of matter. The claim is drawn to a "computer program product stored in a computer readable medium". The specification does not define the computer readable medium, therefore, it includes transitory and non-transitory computer readable medium. The broadest reasonable interpretation of the claim in light of the specification concludes that the claim as a whole covers a transitory signal, which does not fall within the definition of a process, machine, manufacture, or composition of matter (In re Nuijten). Therefore, claim 9 does not fall within a statutory category (the rejection can be overcome by specifying the computer readable medium as "non-transitory computer readable medium").

Claim 14 is rejected under 35 U.S.C. 101 because the claimed invention is directed to non-statutory subject matter. The claim is NOT to a process, machine,

manufacture or composition of matter. The claim is drawn to a "computer program product stored in a computer readable medium". The specification does not define the computer readable medium, therefore, it includes transitory and non-transitory computer readable medium. The broadest reasonable interpretation of the claim in light of the specification concludes that the claim as a whole covers a transitory signal, which does not fall within the definition of a process, machine, manufacture, or composition of matter (In re Nuijten). Therefore, claim 9 does not fall within a statutory category (the rejection can be overcome by specifying the computer readable medium as "non-transitory computer readable medium").

Claim 15-16 are directed to a "video signal"; a signal is not limited to that which falls within a statutory category of invention because it is not limit to a process, machine, manufacture, or a composition of matter. Instead, it includes a form of energy. Energy does not fall with in a statutory category since it is clearly not a series of steps or acts to constitute a process, not a mechanical device or combination of mechanical device to constitute a machine, not a tangible physical article or object which is some form of matter to be a product and constitute a manufacture, and not a composite of two or more substances to constitute a composite of matter.

Claim Rejections - 35 USC § 112

4. The following is a quotation of the first paragraph of 35 U.S.C. 112:

> The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.

Claims 5-8 are rejected under 112, 1st paragraph as being a single component (video encoder) claim, i.e., where a single component (video encoder) recitation does not appear in combination with another recited component, is subject to an undue breadth rejection. The single component which covered every conceivable ways for achieving the stated purpose was held nonenabling for the scope of claim because the specification disclosed at most only those known to the inventor. See MPEP §2164.08(a).

Claim 13 is rejected under 112, 1st paragraph as being a single component (video decoder) claim, i.e., where a single component (video decoder) recitation does not appear in combination with another recited component, is subject to an undue breadth rejection. The single component which covered every conceivable ways for achieving the stated purpose was held nonenabling for the scope of claim because the specification disclosed at most only those known to the inventor. See MPEP §2164.08(a).

Double Patenting

5. The nonstatutory double patenting rejection is based on a judicially created doctrine grounded in public policy (a policy reflected in the statute) so as to prevent the unjustified or improper timewise extension of the "right to exclude" granted by a patent and to prevent possible harassment by multiple assignees. A nonstatutory obviousness-type double patenting rejection is appropriate where the conflicting claims are not identical, but at least one examined application claim is not patentably distinct from the reference claim(s) because the examined application claim is either anticipated by, or would have been obvious over, the reference claim(s). See, e.g., *In re Berg*, 140 F.3d 1428, 46 USPQ2d 1226 (Fed. Cir. 1998); *In re Goodman*, 11 F.3d 1046, 29 USPQ2d 2010 (Fed. Cir. 1993); *In re Longi*, 759 F.2d 887, 225 USPQ 645 (Fed. Cir. 1985); *In re Van Ornum*, 686 F.2d 937, 214 USPQ 761 (CCPA 1982); *In re Vogel*, 422

F.2d 438, 164 USPQ 619 (CCPA 1970); and *In re Thorington*, 418 F.2d 528, 163 USPQ 644 (CCPA 1969).

A timely filed terminal disclaimer in compliance with 37 CFR 1.321(c) or 1.321(d) may be used to overcome an actual or provisional rejection based on a nonstatutory double patenting ground provided the conflicting application or patent either is shown to be commonly owned with this application, or claims an invention made as a result of activities undertaken within the scope of a joint research agreement.

Effective January 1, 1994, a registered attorney or agent of record may sign a terminal disclaimer. A terminal disclaimer signed by the assignee must fully comply with 37 CFR 3.73(b).

6. Claims 1-2, and 4 are provisionally rejected on the ground of nonstatutory obviousness-type double patenting as being unpatentable over claim 9 of copending Application No. 10/306942 in view of Kim et al. (US 2003/0063806 A1, hereinafter Kim).

Claim	Instant application	Claim	Co-pending application 10/306942
1	A method for encoding a video sequence comprising an independent sequence of image frames, wherein at least one reference image frame is predictable from at least one previous image frame that is earlier than the previous reference image frame in decoding order, the	9	Claim 1 A method for coding video frames for the purpose of forming a scalable, compressed video sequence comprising video frames coded according to at least a first and a second frame format, **the video frames of the first frame format being independent of the**

	method comprising: encoding into the video sequence an indication of at least one image frame, which is the first image frame, in decoding order, of the independent sequence, said at least one reference image frame being included in the sequence.		**other video frames**, and **the video frames of the second frame format being predicted from at least one of the other video frames**, the method comprising encoding, at an encoder, the video sequence as at least part of a first sub-sequence,
2	A method according to claim 1, further comprising: encoding the indication into the video sequence as a separate flag included in the header of a slice.		at least part of which has been formed by coding video frames of the at least first frame format; and encoding, at the encoder, the video sequence as at least a
4	A method according to claim 1, further comprising: encoding into the video sequence an identifier value for the independent sequence.		second sub-sequence, at least part of which has been formed by coding at least video frames of the second frame format, and **at least one video frame of the second sub-sequence has been predicted from at least one video frame of the first sub-sequence**; and encoding, at the

			encoder, into the video sequence information indicating which video frames belong to the second sub-sequence. Claim 3 A method according to claim 2, further comprising coding the video sequence into a plural number of scalability layers; and determining the dependency of the video frames of the second sub-sequence **such that at least one video frame of the second sub-sequence is predicted from a group comprising a video frame of a higher scalability layer, a video frame of another sub-sequence in the same scalability layer**. Claim 7 A method according to claim 3, further comprising **determining a**

			unique identifier for each video frame as a combination of layer number, sub-sequence identifier and image number. Claim 9 A method according to claim 7, further comprising **adding the identifier to the header field of the video sequence or to the header field of the transfer protocol to be used for transferring the video sequence**.

This is a provisional obviousness-type double patenting rejection.

(1) Regarding claim 1:

The co-pending application discloses a method for encoding a video sequence comprising an independent sequence of image frames (video sequence comprising video frame coded according to at least a first and second fame format), wherein at least one reference image frame (second sub-sequence) is predictable from at least one previous image frame that is earlier than a previous reference image frame in decoding order (at least one video frame of the second sub-sequence has been

predicted from at least one video frame of the first sub-sequence or a video frame of another sub-sequence in the same scalability layer), the method comprising:

encoding into the video sequence an indication (identifier) of at least one image frame (determined a unique identifier for each video frame), which is the first image frame, in decoding order, of the independent sequence (since **each** frame has an identifier added to the header field of the video sequence, therefore, it included the first image frame), said at least one reference image frame being included in the sequence (the video sequence includes at least a second sub-sequence, therefore, the claimed at least one reference image frame is included in the video sequence).

Claim 9 of the instant application does not explicitly disclose at least one reference image frame is predictable from at least one previous image frame that is earlier than the previous reference image frame in decoding order.

However, Kim discloses an independent video sequence in figure 7 that at least one reference image frame P frame 708) is predictable from at least one previous image frame (I frame 704) that is earlier than a previous reference image frame in decoding order (P frame 706).

It is desirable to have at least one reference image frame is predictable from at least one previous image frame that is earlier than the previous reference image frame in decoding order because it reduces latency when decoding the video sequence. Therefore, it would have been obvious to one of ordinary skill in the art at the time of invention to employ the teaching of Kim in the method of the co-pending application to reduce latency when decoding the video sequence.

(2) Regarding claim 2 (the examiner interprets "a slice" as the video sequence as claim in claim 1):

The copending application discloses:

encoding the indication into the video sequence as a separate flag included in the header of a slice (adding the identifier to the header field of the video sequence or to the header field of the transfer protocol to be used for transferring the video sequence, the examiner interprets the adding process as the encoding process as claim).

(3) Regarding claim 4:

The co-pending application discloses:

encoding into the video sequence an identifier value for the independent sequence (determining a unique identifier for each video frame and adding the identifier to the header field of the video sequence, the examiner interpret adding the unique identifier to the header as encoding the identifier into the video sequence).

7. Claim 3 are provisionally rejected on the ground of nonstatutory obviousness-type double patenting as being unpatentable over claim 9 of copending Application No. 10/306942 in view of Kim et al. (US 2003/0063806 A1, hereinafter Kim) as apply to claim 1 and further in view of Kawashima et al. (US 4,390,966).

Copending application and Kim discloses all subject matter as discussed in claim 1 and copending application further discloses the encoding identifier values for image frames according to a numbering scheme (layer number, sub-sequence identifier and

image number), but fail to disclose resetting the identifier value for the indicated first image frame of an independent sequence.

However, Kawashima discloses a controller that setting or resetting flags as desired according to user instruction (column 6, lines 37-40).

It is desirable to reset the identifier value for the indicated first image frame of an independent sequence because since the video sequence is independent of other sequences, it would simplify the identifier value and reset the identifier value. Therefore, it would have been obvious to one of ordinary skill in the art at the time of invention to employ the resetting step as taught by Kawashima in the method of copending application and Kim to simplify the identifier value of the video sequence.

Kim et al. (US 2003/0063806 A1, hereinafter Kim) in view of Ueda (US 6,094,456)

Claim Rejections - 35 USC § 102

8. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (e) the invention was described in (1) an application for patent, published under section 122(b), by another filed in the United States before the invention by the applicant for patent or (2) a patent granted on an application for patent by another filed in the United States before the invention by the applicant for patent, except that an international application filed under the treaty defined in section 351(a) shall have the effects for purposes of this subsection of an application filed in the United States only if the international application designated the United States and was published under Article 21(2) of such treaty in the English language.

9. Claims 1 and 5 are rejected under 35 U.S.C. 102(e) as being anticipated by Kim et al. (US 2003/0063806 A1, hereinafter Kim).

Kim discloses a method and encoder (encoder discloses in figure 1B and 1C) for encoding a video sequence comprising an independent sequence of image frames (independent sequence as shown in figure 7), wherein at least one reference image frame (as there is no definition of a reference frame, the examiner interprets a P frame as a reference frame, P frame 708 in figure 7) is predictable from at least one previous image frame (I frame 704 in figure 7) that is earlier than the previous reference image frame (P frame 706 in figure 7) in decoding order, the method comprising:

encoding into the video sequence an indication of at least one image frame (a video coding apparatus is provided to receive an indication that a first frame is to be intracoded, paragraph 0018, the examiner interprets the first intracoded frame as the claimed indication), which is the first image frame (the first frame is intracoded), in decoding order, of the independent sequence (independent sequence as shown in figure 7), said at least one reference image frame being included in the sequence (the P frame 708 is included in the sequence).

Claim Rejections - 35 USC § 103

10. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

11. Claims 1, 2, 4, 6, 8, 10, 11, 13, and 15-16 are rejected under 35 U.S.C. 103(a) as being unpatentable over Kim et al. (US 2003/0063806 A1, hereinafter Kim) in view of Ueda (US 6,094,456).

(1) Regarding claim 1:

Kim discloses a method and encoder (encoder discloses in figure 1B and 1C) for encoding a video sequence comprising an independent sequence of image frames (independent sequence as shown in figure 7), wherein at least one reference image frame (as there is no definition of a reference frame, the examiner interprets a P frame as a reference frame, P frame 708 in figure 7) is predictable from at least one previous image frame (I frame 704 in figure 7) that is earlier than the previous reference image frame (P frame 706 in figure 7) in decoding order, the method comprising:

encoding into the video sequence at least one image frame (a video coding apparatus is provided to receive an indication that a first frame is to be intracoded, paragraph 0018, the examiner interprets the first intracoded frame as the claimed indication), which is the first image frame (the first frame is intracoded), in decoding order, of the independent sequence (independent sequence as shown in figure 7), said at least one reference image frame being included in the sequence (the P frame 708 is included in the sequence).

Kim fails to explicitly disclose an indication of the first I frame.

However, Ueda discloses an encoded video sequence that comprises a fixed parameter identifier in the header of an input video code that indicates the type of picture from the picture header (I frame, column 10, lines 38-42).

Since Kim discloses the first frame is an I frame, and Ueda discloses having a fixed parameter identifier in the header to indicate whether the frame is an I frame, it is obvious to employ the teaching of Ueda in the encoding method of Kim to include in the header a fixed parameter identifier for identifying the first I frame in order to improve the reproduction speed (column 2, lines 24-27).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of invention to include in the header a fixed parameter identifier for identifying the first I frame as taught by Ueda in the method of Kim to improve the reproduction speed of the encoded video sequence.

(2) Regarding claims 2 and 6:

Kim discloses all subject matter as discussed in claims 1 and 5 but fail to disclose encoding the indication into the video sequence as a separate flag included in the header of a slice.

However, Ueda discloses an encoded video sequence that comprises a fixed parameter identifier in the header of an input video code that indicates the type of picture from the picture header (column 10, lines 38-42).

Since Kim discloses the first frame is an I frame, and Ueda discloses having a fixed parameter identifier in the header to indicate whether the frame is an I frame, it is obvious to employ the teaching of Ueda in the encoding method of Kim to include in the header a fixed parameter identifier for identifying the first I frame in order to improve the reproduction speed (column 2, lines 24-27).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of invention to include in the header a fixed parameter identifier for identifying the first I frame as taught by Ueda in the method of Kim to improve the reproduction speed of the encoded video sequence.

(3) Regarding claims 4 and 8:

Ueda further discloses encoding into the video sequence an identifier value for the independent sequence (an identifier of six bytes indicative of type of picture, column 7, lines 60-64 and column 10, lines 38-42).

(4) Regarding claims 10 and 13:

Kim discloses a method and decoder for decoding a compressed video sequence (sequence as shown in figure 7), wherein at least one reference image frame (as there is no definition of a reference frame, the examiner interprets a P frame as a reference frame, P frame 708 in figure 7) is predictable from at least one previous image frame (I frame 704 in figure 7) that is earlier than the previous reference image frame (P frame 706 in figure 7) in decoding order, the method comprising

decoding the first two frames of a sequence that yield the exact frame rate, the first two frame as I frame of figure 7 (paragraph 0071) which is the first image frame (I frame 702 in figure 7 is the first frame of the sequence), in decoding order, of an independent sequence, said sequence including the at least one reference image frame (P frame 708 is included in the sequence) and starting the decoding of the video sequence from said first image frame of the independent sequence (as shown in figure 7, the first I frame 702 and the second I frame 704 is for generating the P frame 706-

712, therefore, the decoding processing is start from I frame 702), whereby the video sequence is decoded without prediction from any image frame decoded prior to said first image frame (as shown in figure 7, the sequence is independent and I frame 702 is the first frame of the sequence and no other frame prior to frame 702 is required).

Kim does not explicitly disclose decoding from the video sequence an indication of at least one image frame which is the first image frame, in decoding order of an independent sequence.

However, Ueda discloses a method that determines from an identifier in the header and determine if the frame is an I frame (step 111 and 112 of figure 12, column 14, lines 31-35).

It is desirable to decoding from the video sequence an indication of at least one image frame which is the first image frame, in decoding order of an independent sequence because it improves the reproduction speed (column 2, lines 24-27).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of invention to include in the header a fixed parameter identifier for identifying the first I frame as taught by Ueda in the decoding method of Kim to improve the reproduction speed of the encoded video sequence.

(5) Regarding claim 11:

Ueda further discloses wherein the indication is a separate flag in the header of a slice (an identifier of six bytes indicative of type of picture, column 7, lines 60-64 and column 10, lines 38-42).

(6) Regarding claim 15:

Kim discloses a signal comprising a compressed video sequence including an independent sequence of image frames (sequence as shown in figure 7), wherein at least one reference image frame (as there is no definition of a reference frame, the examiner interprets a P frame as a reference frame, P frame 708 in figure 7) is predictable from at least one previous image frame (I frame 704 in figure 7) that is earlier than the previous reference image frame (P frame 706 in figure 7) in decoding order, the method comprising:

at least one image frame (the first frame is an intracoded frame), which is the first image frame, in decoding order, of the independent sequence (the first I frame 702 is the first frame of the sequence as shown in figure 7), said at least one reference image frame being included in the sequence (P frame 708 is included in the sequence as shown in figure 7).

Kim fails to explicitly disclose an indication of the first I frame.

However, Ueda discloses an encoded video sequence that comprises a fixed parameter identifier in the header of an input video code that indicates the type of picture from the picture header (I frame, column 10, lines 38-42).

Since Kim discloses the first frame is an I frame, and Ueda discloses having a fixed parameter identifier in the header to indicate whether the frame is an I frame, it is obvious to employ the teaching of Ueda in the encoding method of Kim to include in the header a fixed parameter identifier for identifying the first I frame in order to improve the reproduction speed (column 2, lines 24-27).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of invention to include in the header a fixed parameter identifier for identifying the first I frame as taught by Ueda in the method of Kim to improve the reproduction speed of the encoded video sequence.

(7) Regarding claim 16:

Ueda further discloses encoding into the video sequence an identifier value for the independent sequence (an identifier of six bytes indicative of type of picture, column 7, lines 60-64 and column 10, lines 38-42).

12. Claims 3 and 7 are rejected under 35 U.S.C. 103(a) as being unpatentable over Kim et al. (US 2003/0063806 A1, hereinafter Kim) in view of Ueda (US 6,094,456) and Kawashima et al. (US 4,390,966).

Kim discloses all subject matter as discussed in claims 1 and 5 except (a) encoding identifier values for image frames according to a numbering scheme, and (b) resetting the identifier value for the indicated first image frame of an independent sequence.

With respect to (a), Ueda discloses a fixed parameter identifier of six bits in the header of an input video code that indicates the type of picture from the picture header (column 7, lines 60-64 and column 10, lines 38-42, the examiner interprets the bits as a representation of numbering scheme in binary).

Since Kim discloses the first frame is an I frame, and Ueda discloses having a fixed parameter identifier in the header to indicate whether the frame is an I frame, it is

obvious to employ the teaching of Ueda in the encoding method of Kim to include in the header a fixed parameter identifier for identifying the first I frame in order to improve the reproduction speed (column 2, lines 24-27).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of invention to include in the header a fixed parameter identifier for identifying the first I frame as taught by Ueda in the method of Kim to improve the reproduction speed of the encoded video sequence.

With respect to (b), Kawashima discloses a controller that setting or resetting flags as desired according to user instruction (column 6, lines 37-40).

It is desirable to reset the identifier value for the indicated first image frame of an independent sequence because since the video sequence is independent of other sequences, it would simplify the identifier value and reset the identifier value. Therefore, it would have been obvious to one of ordinary skill in the art at the time of invention to employ the resetting step as taught by Kawashima in the method of copending application and Kim to simplify the identifier value of the video sequence.

13. Claims 9 and 14 are rejected under 35 U.S.C. 103(a) as being unpatentable over Kim et al. (US 2003/0063806 A1, hereinafter Kim) in view of Langberg et al. (US 5,852,630).

(1) Regarding claim 9:

Kim discloses a method and encoder (encoder discloses in figure 1B and 1C) for encoding a video sequence comprising an independent sequence of image frames

(independent sequence as shown in figure 7), wherein at least one reference image frame (as there is no definition of a reference frame, the examiner interprets a P frame as a reference frame, P frame 708 in figure 7) is predictable from at least one previous image frame (I frame 704 in figure 7) that is earlier than the previous reference image frame (P frame 706 in figure 7) in decoding order, the method comprising:

encoding into the video sequence an indication of at least one image frame (a video coding apparatus is provided to receive an indication that a first frame is to be intracoded, paragraph 0018, the examiner interprets the first intracoded frame as the claimed indication), which is the first image frame (the first frame is intracoded), in decoding order, of the independent sequence (independent sequence as shown in figure 7), said at least one reference image frame being included in the sequence (the P frame 708 is included in the sequence).

Kim fails to disclose implementing the encoder by a computer program product, stored on a computer readable medium and executable in a data processing unit.

However, Langberg et al. teaches that the method and apparatus for a transceiver warm start activation procedure with precoding can be implemented in software stored in a computer-readable medium. The computer-readable medium is an electronic, magnetic, optical, or other physical device or means that can be contain or store a computer program for use by or in connection with a computer-related system or method (column 3, lines 51-65). One skilled in the art would have clearly recognized that the method of Kim would have been implemented in a software. The implemented software would perform same function of the hardware for less expense, adaptability,

and flexibility. Therefore, it would have been obvious to have used the software as taught by Langberg to implement the method and encoder of Kim in order to reduce cost and improve the adaptability and flexibility of the communication system.

(2) Regarding claim 14:

Kim discloses a method and decoder for decoding a compressed video sequence (sequence as shown in figure 7), wherein at least one reference image frame (as there is no definition of a reference frame, the examiner interprets a P frame as a reference frame, P frame 708 in figure 7) is predictable from at least one previous image frame (I frame 704 in figure 7) that is earlier than the previous reference image frame (P frame 706 in figure 7) in decoding order, the method comprising

decoding the first two frames of a sequence that yield the exact frame rate, the first two frame as I frame of figure 7 (paragraph 0071) which is the first image frame (I frame 702 in figure 7 is the first frame of the sequence), in decoding order, of an independent sequence, said sequence including the at least one reference image frame (P frame 708 is included in the sequence) and starting the decoding of the video sequence from said first image frame of the independent sequence (as shown in figure 7, the first I frame 702 and the second I frame 704 is for generating the P frame 706-712, therefore, the decoding processing is start from I frame 702), whereby the video sequence is decoded without prediction from any image frame decoded prior to said first image frame (as shown in figure 7, the sequence is independent and I frame 702 is the first frame of the sequence and no other frame prior to frame 702 is required).

Kim does not explicitly disclose (a) decoding from the video sequence an indication of at least one image frame which is the first image frame, in decoding order of an independent sequence, and (b) implementing the method by a computer program product, stored on a computer readable medium and executable in a data processing unit.

With respect to (a), Ueda discloses a method that determines from an identifier in the header and determine if the frame is a I frame (step 111 and 112 of figure 12, column 14, lines 31-35).

It is desirable to decoding from the video sequence an indication of at least one image frame which is the first image frame, in decoding order of an independent sequence because it improves the reproduction speed (column 2, lines 24-27).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of invention to include in the header a fixed parameter identifier for identifying the first I frame as taught by Ueda in the decoding method of Kim to improve the reproduction speed of the encoded video sequence.

With respect to (b), Langberg et al. teaches that the method and apparatus for a transceiver warm start activation procedure with precoding can be implemented in software stored in a computer-readable medium. The computer-readable medium is an electronic, magnetic, optical, or other physical device or means that can be contain or store a computer program for use by or in connection with a computer-related system or method (column 3, lines 51-65). One skilled in the art would have clearly recognized that the method of Kim would have been implemented in a software. The implemented

software would perform same function of the hardware for less expense, adaptability, and flexibility. Therefore, it would have been obvious to have used the software as taught by Langberg to implement the method and encoder of Kim and Ueda in order to reduce cost and improve the adaptability and flexibility of the communication system.

14. Claim 12 is rejected under 35 U.S.C. 103(a) as being unpatentable over Kim et al. (US 2003/0063806 A1, hereinafter Kim) in view of Ueda (US 6,094,456) as applied to claim 10 and further in vie of Kawashima et al. (US 4,390,966).

Kim and Ueda disclose all subject matter as discussed in claim 10, and Ueda further discloses a fixed parameter identifier of six bits in the header of an input video code that indicates the type of picture from the picture header (column 7, lines 60-64 and column 10, lines 38-42, the examiner interprets the bits as a representation of numbering scheme in binary).

Kim and Ueda fail to disclose resetting the identifier value for the indicated first image frame of an independent sequence.

However, Kawashima discloses a controller that setting or resetting flags as desired according to user instruction (column 6, lines 37-40).

It is desirable to reset the identifier value for the indicated first image frame of an independent sequence because since the video sequence is independent of other sequences, it would simplify the identifier value and reset the identifier value. Therefore, it would have been obvious to one of ordinary skill in the art at the time of

invention to employ the resetting step as taught by Kawashima in the method of copending application and Kim to simplify the identifier value of the video sequence.

Conclusion

15. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

Cuccia (US 202/0051621 A1) discloses an instant replay on digital video optimized using non MPEG frame tag.

Hannuksela et al. (US 2001/0040700 A1) discloses a video coding.

Apostolopoulos (US 7,103,669 B2) discloses a video communication method and system employing multiple state encoding and path diversity.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to SIU M. LEE whose telephone number is (571)270-1083. The examiner can normally be reached on Mon-Fri, 7:30-4:00 with every other Friday off.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Chieh Fan can be reached on (571) 272-3042. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Siu M Lee/
Examiner, Art Unit 2611
5/21/2010

/CHIEH M FAN/

Supervisory Patent Examiner, Art Unit 2611

Notice of References Cited	Application/Control No. 11/338,934	Applicant(s)/Patent Under Reexamination HANNUKSELA, MISKA	
	Examiner SIU M. LEE	Art Unit 2611	Page 1 of 1

U.S. PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Name	Classification
*	A	US-6,094,456	07-2000	Ueda, Hiroaki	375/240.12
*	B	US-2003/0063806	04-2003	Kim et al.	382/236
*	C	US-4,390,966	06-1983	Kawashima et al.	712/229
*	D	US-7,103,669	09-2006	Apostolopoulos, John G.	709/231
*	E	US-2001/0040700	11-2001	Hannuksela et al.	358/261.2
*	F	US-2002/0051621	05-2002	Cuccia, David W.	386/68
	G	US-			
	H	US-			
	I	US-			
	J	US-			
	K	US-			
	L	US-			
	M	US-			

FOREIGN PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Country	Name	Classification
	N					
	O					
	P					
	Q					
	R					
	S					
	T					

NON-PATENT DOCUMENTS

*		Include as applicable: Author, Title Date, Publisher, Edition or Volume, Pertinent Pages)
	U	
	V	
	W	
	X	

*A copy of this reference is not being furnished with this Office action. (See MPEP § 707.05(a).)
Dates in MM-YYYY format are publication dates. Classifications may be US or foreign.

Index of Claims	Application/Control No.	Applicant(s)/Patent Under Reexamination
	11338934	HANNUKSELA, MISKA
	Examiner	Art Unit
	SIU M LEE	2611

✓	Rejected	-	Cancelled	N	Non-Elected	A	Appeal
=	Allowed	÷	Restricted	I	Interference	O	Objected

☐ Claims renumbered in the same order as presented by applicant ☐ CPA ☐ T.D. ☐ R.1.47

CLAIM		DATE								
Final	Original	05/21/2010								
	1	✓								
	2	✓								
	3	✓								
	4	✓								
	5	✓								
	6	✓								
	7	✓								
	8	✓								
	9	✓								
	10	✓								
	11	✓								
	12	✓								
	13	✓								
	14	✓								
	15	✓								
	16	✓								

Search Notes	Application/Control No.	Applicant(s)/Patent Under Reexamination
	11338934	HANNUKSELA, MISKA
	Examiner	Art Unit
	SIU M LEE	2611

SEARCHED

Class	Subclass	Date	Examiner
375	240.12, 240.25	5/21/2010	Siu M. Lee

SEARCH NOTES

Search Notes	Date	Examiner
EAST	5/21/2010	Siu M. Lee
Inventor name search - Miska hannuksela	5/21/2010	Siu M. Lee
Discuss the application with Chieh Fan	5/21/2010	Siu M. Lee

INTERFERENCE SEARCH

Class	Subclass	Date	Examiner

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

BIB DATA SHEET

CONFIRMATION NO. 6610

SERIAL NUMBER	FILING or 371(c) DATE RULE	CLASS	GROUP ART UNIT	ATTORNEY DOCKET NO.
11/338,934	01/25/2006	375	2611	TPO.004.C2

APPLICANTS
Miska Hannuksela, Tampere, FINLAND;

**** CONTINUING DATA ***************************
This application is a CON of 10/306,942 11/29/2002

**** FOREIGN APPLICATIONS ***************************
FINLAND 20020127 01/23/2002

**** IF REQUIRED, FOREIGN FILING LICENSE GRANTED ****
02/28/2006

Foreign Priority claimed ☑ Yes ☐ No 35 USC 119(a-d) conditions met ☑ Yes ☐ No Verified and Acknowledged /SIU M LEE/ Examiner's Signature	☐ Met after Allowance Initials	STATE OR COUNTRY	SHEETS DRAWINGS	TOTAL CLAIMS	INDEPENDENT CLAIMS
		FINLAND	4	16	7

ADDRESS
Hollingsworth & Funk
8500 Normandale Lake Blvd., Suite 320
Minneapolis, MN 55437
UNITED STATES

TITLE
Grouping of image frames in video coding

FILING FEE RECEIVED	FEES: Authority has been given in Paper	☐ All Fees
1800	No._____ to charge/credit DEPOSIT ACCOUNT	☐ 1.16 Fees (Filing)
	No._____ for following:	☐ 1.17 Fees (Processing Ext. of time)
		☐ 1.18 Fees (Issue)
		☐ Other _____
		☐ Credit

EAST Search History

EAST Search History (Prior Art)

Ref #	Hits	Search Query	DBs	Default Operator	Plurals	Time Stamp
L1	4706	375/240.12,240.25.ccls.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/21 07:33
L2	2292	1 and (@pd<="20020123" OR @ad<="20020123" OR @rlad<="20020123" OR @prad<="20020123")	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/21 07:33
L3	39	2 and (video near2 encod$3) and (first frame with (indicat$3 or tag or identifier or flag))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/21 07:35
L4	2	"6094456".pn. and identifier	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/21 09:19
L5	12	10/092340 and decod$3	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/21 10:11
S1	1	11/338934 and program	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/19 14:20
S2	190	(miska near2 Hannuksela).in.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/19 15:30

S3	149	S2 and stream$3	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/19 15:30
S4	38	S3 and (@pd<="20020123" OR @ad<="20020123" OR @rlad<="20020123" OR @prad<="20020123")	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/19 15:31
S5	2	"7289506".pn.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/19 15:33
S6	1	09/560217	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/19 15:33
S7	21030	encod$3 with (picture or frame) with (indicat$3 or tag or flag or reference)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 09:17
S8	12392	S7 and (@pd<="20020123" OR @ad<="20020123" OR @rlad<="20020123" OR @prad<="20020123")	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 09:17
S9	567	S8 and (encod$3 with (picture or frame) with ((indicat$3 or tag or flag or reference) near3 first))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 09:18
S10	43406	scalability	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 09:19
S11	45	S9 and S10	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 09:19

S12	2	"7236526".pn.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 09:21
S13	7243	encod$3 with (picture or frame) with (indicator or tag or flag or label$4)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 09:41
S14	4794	S13 and (@pd<="20020123" OR @ad<="20020123" OR @rlad<="20020123" OR @prad<="20020123")	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 09:41
S15	149	S14 and scalability	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 09:41
S16	12	S15 and ((indicat$3 or tag or flag or reference) near3 (I or independent or first) frame)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 09:42
S30	275	(indicator or flag or identifier or tag or bit) with header with (i or independent or first) frame	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 11:52
S31	136	S30 and (@pd<="20020123" OR @ad<="20020123" OR @rlad<="20020123" OR @prad<="20020123")	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 11:52
S32	331	(indicator or flag or identifier or tag or bit) with header with (i or independent or first or intra) frame	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 11:52
S33	177	S32 and (@pd<="20020123" OR @ad<="20020123" OR @rlad<="20020123" OR @prad<="20020123")	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 11:52

S34	128	S33 and video	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 11:53
S35	2	"7336787".pn.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 12:07
S36	1	11/338934 and (medium)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 13:31
S37	1	11/338934 and (computer readable medium)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 13:31
S38	751	identifier with flag with header	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 14:47
S39	1	identifier with flag with header with video sequence	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 14:47
S40	4	identifier with flag with header with start with video	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 14:48
S41	20	encod$3 with identifier with flag with header	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 14:49
S42	2	resetting with (identifier or tag or indicator or flag) with independent with sequence	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 15:09

S43	9	video sequence with (identifier or tag or indicator or flag) with independent with sequence	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 15:36
S44	2	"6968005".pn.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 15:39

EAST Search History (Interference)

Ref #	Hits	Search Query	DBs	Default Operator	Plurals	Time Stamp
S45	140	(miska near2 Hannuksela).in.	US-PGPUB; USPAT; UPAD	ADJ	ON	2010/05/20 06:40
S46	139	S45 and stream$3	US-PGPUB; USPAT; UPAD	ADJ	ON	2010/05/20 06:41
S47	37	S46 and (@pd<="20020123" OR @ad<="20020123" OR @rlad<="20020123" OR @prad<="20020123")	US-PGPUB; USPAT; UPAD	ADJ	ON	2010/05/20 06:41
S48	37	S47 not "s4"	US-PGPUB; USPAT; UPAD	ADJ	ON	2010/05/20 06:41
S49	18599	encod$3 with (picture or frame) with (indicat$3 or tag or flag or reference)	US-PGPUB; USPAT; UPAD	ADJ	ON	2010/05/20 09:17

5/21/2010 11:49:33 AM
C:\Documents and Settings\slee7\My Documents\EAST\Workspaces\11338934 201005.wsp

FORM 1449* INFORMATION DISCLOSURE STATEMENT IN AN APPLICATION (Use several sheets if necessary)	Docket Number: TPO.004.C2	Application Number: 11/338,934
	Applicant: HANNUKSELA	
	Filing Date: 1/25/2006	Group Art Unit: 2613

U.S. PATENT DOCUMENTS

EXAMINER INITIAL	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE

FOREIGN PATENT DOCUMENTS

	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	TRANSLATION	
						YES	NO
/SL/	11-177987	02.07.1999	JP			Abstract	

OTHER DOCUMENTS (Including Author, Title, Date, Pertinent Pages, Etc.)

/SL/	2010	English Translation of the Office Action of parallel Japanese application no. 2006-120296 dated 01/06/2010.

EXAMINER /Siu Lee/	DATE CONSIDERED 05/20/2010

EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609; draw line through citation if not in conformance and not considered. Include copy of this form for next communication to the Applicant.

FORM 1449* INFORMATION DISCLOSURE STATEMENT IN AN APPLICATION (Use several sheets if necessary)	Docket Number: TPO.004.C2	Application Number: 11/338,934
	Applicant: HANNUKSELA	
	Filing Date: Jan. 25, 2006	Group Art Unit: 2613

OIPE IAP76 JUN 16 2008 PATENT & TRADEMARK OFFICE

U.S. PATENT DOCUMENTS

EXAMINER INITIAL	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
/SL/	20010024472	09/27/2001	Sporer et al.			

FOREIGN PATENT DOCUMENTS

	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	TRANSLATION	
						YES	NO

OTHER DOCUMENTS (Including Author, Title, Date, Pertinent Pages, Etc.)

EXAMINER /Siu Lee/	DATE CONSIDERED 05/20/2010

EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609; draw line through citation if not in conformance and not considered. Include copy of this form for next communication to the Applicant.

*Substitute Disclosure Statement Form (PTO-1449)

FORM 1449* INFORMATION DISCLOSURE STATEMENT IN AN APPLICATION (Use several sheets if necessary)	Docket Number: TPO.004.C2	Application Number: 11/338,934
	Applicant: HANNUKSELA	
	Filing Date: 1/25/2006	Group Art Unit: 2613

U.S. PATENT DOCUMENTS

EXAMINER INITIAL	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
/SL/	6496980	12/17/2002	Tillman et al.			
/SL/	6510553	01/21/2003	Hazra			
/SL/	6639943	10/28/2003	Radha et al.			
/SL/	6614936	09/02/2003	Wu et al.			

FOREIGN PATENT DOCUMENTS

	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	TRANSLATION	
						YES	NO
/SL/	EP0898427	2/24/1999	EP			NA	

OTHER DOCUMENTS (Including Author, Title, Date, Pertinent Pages, Etc.)

/SL/	5/7/1996	Elloumi et al., "Issues in Multimedia Scaling: the MPEG Video Streams Case", Emerging Technologies and Applications in Communications, IEEE Computer Society, May 7, 1996.
/SL/	2/1/2002	Hannuksela, "Enhanced Concept of GOP", Joint Video Team of ISO-IES MPEG 7 itu-T VCEG, 2/1/2002.
/SL/	10/12/1997	Soderquist et al., "Memory Traffic and Data Cache Behavior of an MPEG-2 Software Decoder", IEEE Computer Society, October 12, 1997.
/SL/	2/2/2010	Copy of Office Action of related European application no. 06101013.8 dated 02/02/2010.
/SL/	2/2/2010	Copy of Office Action of related European application no. 06101014.6 dated 02/02/2010.
/SL/	3/16/2006	Copy of Office Action of parallel U.S. Application No. 10/306,942 dated 3/16/2006, 4 pages.
/SL/	8/10/2006	Copy of Office Action of parallel U.S. Application No. 10/306,942 dated 8/10/2006, 16 pages.
/SL/	9/11/2007	Copy of Office Action of parallel U.S. Application No. 10/306,942 dated 9/11/2007, 4 pages.
/SL/	3/13/2007	Copy of Office Action of parallel U.S. Application No. 10/306,942 dated 3/13/2007, 7 pages.
/SL/	12/7/2009	Copy of Notice of Allowance of parallel U.S. Application No. 10/306,942 dated 12/7/2009, 8 pages.

EXAMINER /Siu Lee/	DATE CONSIDERED 05/20/2010

EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609; draw line through citation if not in conformance and not considered. Include copy of this form for next communication to the Applicant.

*Substitute Disclosure Statement Form (PTO-1449) Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

FORM 1449* INFORMATION DISCLOSURE STATEMENT IN AN APPLICATION (Use several sheets if necessary)	Docket Number: TPO.004.C2	Application Number: 11/338,934
	Applicant: HANNUKSELA	
	Filing Date: 1/25/2006	Group Art Unit: 2613

/SL/	5/6/2006	Copy of Office Action Response of parallel U.S. Application No. 10/306,942 dated 8/6/2006, 13 pages.
/SL/	11/9/2006	Copy of Office Action Response of parallel U.S. Application No. 10/306,942 dated 11/9/2006, 16 pages.
/SL/	5/31/2007	Copy of Office Action Response of parallel U.S. Application No. 10/306,942 dated 5/31/2007, 12 pages.
/SL/	7/9/2009	Copy of Office Action Response of parallel U.S. Application No. 10/306,942 dated 7/9/2009, 12 pages.

EXAMINER /Siu Lee/	DATE CONSIDERED 05/20/2010
EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609; draw line through citation if not in conformance and not considered. Include copy of this form for next communication to the Applicant.	

FORM 1449* INFORMATION DISCLOSURE STATEMENT IN AN APPLICATION (Use several sheets if necessary)	Docket Number: TPO.004.C2	Application Number: 11/338,934
	Applicant: HANNUKSELA	
	Filing Date: Jan. 25, 2006	Group Art Unit: 2613

U.S. PATENT DOCUMENTS

EXAMINER INITIAL	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
/SL/	6,26,158	7/24/2001	Hata et al.			

FOREIGN PATENT DOCUMENTS

	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	TRANSLATION	
						YES	NO
/SL/	10-271511	10/09/1998	JP				X

OTHER DOCUMENTS (Including Author, Title, Date, Pertinent Pages, Etc.)

EXAMINER /Siu Lee/	DATE CONSIDERED 05/20/2010

EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609; draw line through citation if not in conformance and not considered. Include copy of this form for next communication to the Applicant.

*Substitute Disclosure Statement Form (PTO-1449)

FORM 1449* SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT IN AN APPLICATION	Docket Number: TAPA.004.C2	Application Number: 11/338,934
	Applicant: HANNUKSELA	
	Filing Date: Jan. 25, 2006	Group Art Unit: 2613

OIPE IAP17 OCT 22 2007 PATENT & TRADEMARK OFFICE

U.S. PATENT DOCUMENTS

EXAMINER INITIAL	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
/SL/	5,144,426	09.01.1992	Tanaka et al.			

FOREIGN PATENT DOCUMENTS

	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	TRANSLATION	
						YES	NO
/SL/	01/84850	8 Nov 2001	WO			NA	

OTHER DOCUMENTS (Including Author, Title, Date, Pertinent Pages, Etc.)

EXAMINER	/Siu Lee/	DATE CONSIDERED	05/20/2010

EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609; draw line through citation if not in conformance and not considered. Include copy of this form for next communication to the Applicant.

*Substitute Disclosure Statement Form (PTO-1449) Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

FORM 1449* INFORMATION DISCLOSURE STATEMENT IN AN APPLICATION (Use several sheets if necessary)	Docket Number: TPO.004.C2	Application Number: 11/338,934
	Applicant: HANNUKSELA	
	Filing Date: 1/25/2006	Group Art Unit: 2613

U.S. PATENT DOCUMENTS

EXAMINER INITIAL	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
/SL/	5122875	06/16/1992	Raychaudhuri et al.			
/SL/	5838265	11/17/1998	Adolph			
/SL/	5877812	03/02/1999	Krause et al.			
/SL/	6314139	11/06/2001	Koto et al.			

FOREIGN PATENT DOCUMENTS

	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	TRANSLATION	
						YES	NO

OTHER DOCUMENTS (Including Author, Title, Date, Pertinent Pages, Etc.)

/SL/	02/04/2010	Copy of Allowance Notification with translation dated 2/4/2010 from parallel Russian Application No. 2006110322, 17 pages.
/SL/	02/25/2010	Copy of translated Office Action dated 2/25/2010 from parallel Japanese Application No. 2006-120297, 5 pages.

EXAMINER /Siu Lee/	DATE CONSIDERED 05/20/2010

EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609; draw line through citation if not in conformance and not considered. Include copy of this form for next communication to the Applicant.

*Substitute Disclosure Statement Form (PTO-1449)

FORM 1449* INFORMATION DISCLOSURE STATEMENT IN AN APPLICATION (Use several sheets if necessary)	Docket Number: TPO.004.C2	Application Number: 11/338,934
	Applicant: HANNUKSELA	
	Filing Date: 01/25/2006	Group Art Unit: 2613

OIPE IAP2 JAN 19 2006 PATENT & TRADEMARK OFFICE

U.S. PATENT DOCUMENTS

EXAMINER INITIAL	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
/SL/	5,699,476	16.12.1997	VAN DER MEER			
	5,774,593	30.6.1998	ZICK ET AL.			
	6,108,382	22.8.2000	GRINGERI ET AL.			
	6,307,886	23.10.2001	WESTERMANN			
/SL/	6,337,881	8.1.2002	CHADDHA			

FOREIGN PATENT DOCUMENTS

	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	TRANSLATION	
						YES	NO
/SL/	WO 00/67469	9.11.2000	PCT			n/a	
	WO 01/47283	28.06.2001	PCT			n/a	
	WO 02/054776	11.7.2002	PCT			n/a	
/SL/	EP 1 005 232	31.05.2000	EP			n/a	

OTHER DOCUMENTS (Including Author, Title, Date, Pertinent Pages, Etc.)

/SL/	9-17 Oct. 2002	Joint Video Team of ISO/IEC MPEG and ITU-T VCEG, *Editor's Proposed Draft Text Modification for Joint Video Specification (ITU-T Rec. 264 ISO/IEC 14496-10 AVC)*, Geneva modifications draft 37.
/SL/	15 Jan. 1998	ITU-Telecommunications Standardization Union, *Draft Text of Recommendation H.263 Version 2 ("H.263+") for Decision.*
/SL/	7/10/02	ITU-Telecommunications Standardization Union, *H.26L Test Model Long Term Number 8 (TML-8) Draft0*, Document VCEG-N10. pp. 1-46.
/SL/	Nov. 2000	ITU-Telecommunications Standardization Union, *Draft for "H.263++" Annexes U,V, and W to Recommendation H.263.*
/SL/	12 Nov. 2002	David Singer and Toby Walker, *Study Text of ISO/IEC 14496-15/FCD,* pp. 1-25.
/SL/	12 Nov. 2002	David Singer and Toby Walker, *Study Text of ISO/IEC 14496-15/FCD (Revised), 26* pages.
/SL/	Sept. 2002	Hannuksela, M et al.; Sub-Picture: ROI Coding and Unequal Error Protection, IEEE 2002 Conference, September 2002, USA (internet publication).
/SL/	Sept. 2002	Hannuksela, M et al.; Sub-Picture: Video-Coding for Unequal Error Protection, XI European Signal Processing Conference, September 2002, Toulouse, France (internet publication).
/SL/	7.12.2001	Wang, Y.K. et al.; Core Experiment Description of Sub-Picture Coding, ITU 15. Meeting 3-7.12.2001, Pattaya, Thailand (internet publication)
/SL/	2001	B.A. Lokshin, "Digital Broadcasting: from Studio to Audience" Cyrus Systems, Moscow, 2001. (translation included)

EXAMINER /Siu Lee/	DATE CONSIDERED 05/20/2010

EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609; draw line through citation if not in conformance and not considered. Include copy of this form for next communication to the Applicant.

FORM 1449* INFORMATION DISCLOSURE STATEMENT IN AN APPLICATION (Use several sheets if necessary)	Docket Number: TPO.004.C2	Application Number: 11/338,934
	Applicant: HANNUKSELA	
	Filing Date: Jan. 25, 2006	Group Art Unit: 2613

U.S. PATENT DOCUMENTS

EXAMINER INITIAL	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE

FOREIGN PATENT DOCUMENTS

	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	TRANSLATION	
						YES	NO
/SL/	EP 0798932	10/1/1997	EP			NA	
/SL/	CN 1173783A	2/18/1998	China				X

OTHER DOCUMENTS (Including Author, Title, Date, Pertinent Pages, Etc.)

EXAMINER /Siu Lee/	DATE CONSIDERED 05/20/2010

EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609; draw line through citation if not in conformance and not considered. Include copy of this form for next communication to the Applicant.

FORM 1449* INFORMATION DISCLOSURE STATEMENT IN AN APPLICATION (Use several sheets if necessary)	Docket Number: TPO.004.C2	Application Number: 11/338,934
	Applicant: HANNUKSELA	
	Filing Date: 1/25/2006	Group Art Unit: 2613

OIPE IAP19 AUG 18 2006 PATENT & TRADEMARK OFFICE

U.S. PATENT DOCUMENTS

EXAMINER INITIAL	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE

FOREIGN PATENT DOCUMENTS

	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	TRANSLATION	
						YES	NO

OTHER DOCUMENTS (Including Author, Title, Date, Pertinent Pages, Etc.)

/SL/	1998	Wenger, "Temporal Scalability Using P-pictures for Low-latency Applications", IEEE Workshop, December 1998, pp. 559-564.

EXAMINER /Siu Lee/	DATE CONSIDERED 05/20/2010

EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609; draw line through citation if not in conformance and not considered. Include copy of this form for next communication to the Applicant.

*Substitute Disclosure Statement Form (PTO-1449) Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

IAP43
NOV 24 2006
PATENT & TRADEMARK OFFICE

FORM 1449*	INFORMATION DISCLOSURE STATEMENT	Docket Number: TPO.004.C2	Application Number: 11/338,934
	IN AN APPLICATION	Applicant: HANNUKSELA	
	(Use several sheets if necessary)	Filing Date: 1/25/2006	Group Art Unit: 2613

U.S. PATENT DOCUMENTS

EXAMINER INITIAL	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE

FOREIGN PATENT DOCUMENTS

	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	TRANSLATION	
						YES	NO
/SL/	RU 2 123 769	20.12.1998	Russia			Abstract Only	

OTHER DOCUMENTS (Including Author, Title, Date, Pertinent Pages, Etc.)

EXAMINER	/Siu Lee/	DATE CONSIDERED	05/20/2010

EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609; draw line through citation if not in conformance and not considered. Include copy of this form for next communication to the Applicant.

*Substitute Disclosure Statement Form (PTO-1449) Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

FORM 1449* INFORMATION DISCLOSURE STATEMENT IN AN APPLICATION (Use several sheets if necessary)	Docket Number: TPO.004.C2	Application Number: 11/338,934
	Applicant: HANNUKSELA	
	Filing Date: 1/25/2006	Group Art Unit: 2611

U.S. PATENT DOCUMENTS

EXAMINER INITIAL	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	5852630	12/22/1998	Langberg			
	6072831	06/06/2000	Chen			
	6363208	03/26/2002	Nitta et al.			

FOREIGN PATENT DOCUMENTS

	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	TRANSLATION	
						YES	NO
	2137197C1	9/10/1999	Russia				

OTHER DOCUMENTS (Including Author, Title, Date, Pertinent Pages, Etc.)

	1997	Illgner et al., "Spatially Scalable Video Compression Employing Resolution Pyramids", IEEE Journal on Selected Areas in Communications, Vol. 15, No. 9, December 1, 2007, pg. 1688-1703.
	2000	Information Technology – Generic Coding of Moving Pictures and Associated Audio Information: Systems, ISO/IEC 13818-1, Second edition, December 1, 2000, 174 pages.
	5/5/2010	Notice of Allowance dated 5/5/2010 from Russian Application No. 2006110321/09, 11 pages.
	6/17/2010	Office Action dated 6/17/2010 from U.S. Serial No. 11/338,996, 30 pages.
	6/23/2010	Office Action dated 6/23/2010 from U.S. Serial No. 10/306,942, 6 pages.
	5/25/2010	Office Action dated 5/25/2010 from U.S. Serial No. 10/306,942, 6 pages.
	4/09/2009	Office Action dated 4/9/2009 from U.S. Serial No. 10/306,942, 4 pages.
	5/8/2006	Office Action Response dated 5/8/2006 from U.S. Serial No. 10/306,942, 11 pages.

EXAMINER	DATE CONSIDERED
EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609; draw line through citation if not in conformance and not considered. Include copy of this form for next communication to the Applicant.	

*Substitute Disclosure Statement Form (PTO-1449) Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

(19) **RU** (11) **2 137 194** (13) **C1**

(51) МПК[6] **G 06 K 9/00, G 06 T 7/20**

РОССИЙСКОЕ АГЕНТСТВО
ПО ПАТЕНТАМ И ТОВАРНЫМ ЗНАКАМ

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ РОССИЙСКОЙ ФЕДЕРАЦИИ

(21), (22) Заявка: 98114418/09, 15.07.1998

(24) Дата начала действия патента: 15.07.1998

(46) Дата публикации: 10.09.1999

(56) Ссылки: K.R.RAO, J.J.HWANG, "Techniques and Standards for Image, Video and Audio Coding", 1996, PRENTICE-HALL PTR, ISBN 0-13-309907-5. SU 453978 A, 07.03.84. SU 802974 A, 07.02.81. EP 0647919 A1, 12.04.95. GB 2282021 A, 22.03.95.

(98) Адрес для переписки:
103064, Москва, ул.Казакова, 16,
НИИР-Канцелярия "Патентные поверенные Квашнин, Сапельников и Партнеры" Патентному поверенному Сапельникову Д.А.

(71) Заявитель:
Дворкович Александр Викторович,
Дворкович Виктор Павлович,
Зубарев Юрий Борисович,
Соколов Александр Юрьевич

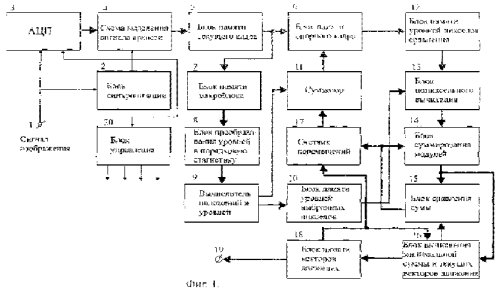
(72) Изобретатель: Дворкович А.В.,
Дворкович В.П., Зубарев Ю.Б., Соколов А.Ю.

(73) Патентообладатель:
Дворкович Александр Викторович,
Дворкович Виктор Павлович,
Зубарев Юрий Борисович,
Соколов Александр Юрьевич

(54) СПОСОБ АНАЛИЗА ВЕКТОРОВ ДВИЖЕНИЯ ДЕТАЛЕЙ В ДИНАМИЧЕСКИХ ИЗОБРАЖЕНИЯХ

(57) Реферат:

Изобретение относится к видеоинформационной технике и может найти применение при разработке цифровых кодирующих устройств для видеотелефонии, видеоконференцсвязи, телевизионного цифрового вещания стандартной и высокой четкости. Технический результат заключается в сокращении количества вычислительных операций, увеличении зоны поиска векторов движения деталей в динамических изображениях и, вследствие этого, повышении качества воспроизведения быстро перемещающихся деталей. Технический результат достигается за счет того, что в способе анализа векторов движения деталей в динамических изображениях все множество пикселов каждого из макроблоков изображения разбивается на участки, в каждом из которых выбирается только один пиксел таким образом, что уровни всей совокупности выбранных пикселов характеризуют изменение рельефа сигнала яркости макроблока текущего кадра, и затем контрольная сумма, являющаяся суммой норм попиксельной разности уровней в текущем и опорном кадрах, вычисляется лишь по выбранным упомянутым пикселам. 1 з.п. ф-лы, 8 ил., 4 табл.



Фиг. 1

RU 2137194 C1

RU 2137194 C1

(19) **RU** (11) **2 137 194** (13) **C1**

(51) Int. Cl.[6] **G 06 K 9/00, G 06 T 7/20**

RUSSIAN AGENCY
FOR PATENTS AND TRADEMARKS

(12) ABSTRACT OF INVENTION

(21), (22) Application: **98114418/09, 15.07.1998**

(24) Effective date for property rights: **15.07.1998**

(46) Date of publication: **10.09.1999**

(98) Mail address:
103064, Moskva, ul.Kazakova, 16, NIIR-Kantseljarija "Patentnye poverennye Kvashnin, Sapel'nikov i Partnery" Patentnomu poverennomu Sapel'nikovu D.A.

(71) Applicant:
Dvorkovich Aleksandr Viktorovich,
Dvorkovich Viktor Pavlovich,
Zubarev Jurij Borisovich,
Sokolov Aleksandr Jur'evich

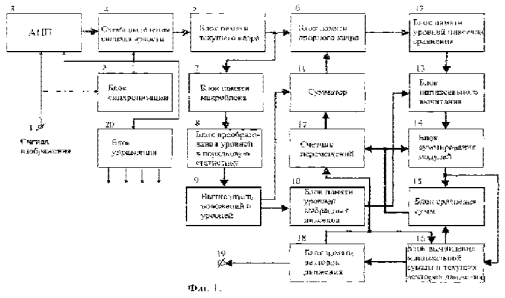
(72) Inventor: **Dvorkovich A.V.,**
Dvorkovich V.P., Zubarev Ju.B., Sokolov A.Ju.

(73) Proprietor:
Dvorkovich Aleksandr Viktorovich,
Dvorkovich Viktor Pavlovich,
Zubarev Jurij Borisovich,
Sokolov Aleksandr Jur'evich

(54) METHOD FOR ANALYSIS OF MOVEMENT VECTORS OF ELEMENTS IN DYNAMIC IMAGES

(57) Abstract:

FIELD: video presentation devices, in particular, digital encoders for video telephones, conference calls, standard and high-definition TV systems. SUBSTANCE: method involves splitting complete set of pixels of each large image block into regions in which only one pixel is chosen so that levels of complete set of chosen pixels define alteration of shape of brightness signal of large image block of current frame. Then check sum which is defined as sum of norms of by-pixel difference in levels in current and reference frames is calculated using data from only chosen pixels. EFFECT: decreased amount of calculations, increased region for search of movement vectors of elements, improved quality of representation of fast moving elements. 2 cl, 13 dwg, 4 tbl



RU 2137194 C1

RU 2137194 C1

Изобретение относится к видеоинформационной технике и может найти применение при разработке цифровых кодирующих устройств для видеотелефонии, видеоконференцсвязи, телевизионного цифрового вещания стандартной и высокой четкости.

Известно множество способов анализа векторов движения деталей в динамических изображениях.

Наиболее простым и точным является способ поиска векторов движения макроблоков на основе полного перебора [1, с. 89-91]. Для поиска вектора движения $\vec{V} = (V_x, V_y)$ рассматривается норма разницы сигналов яркости двух макроблоков в текущем и опорном кадрах SAD со сдвигом на вектор движения:

$$SAD = \sum_{x,y=1,16} |F(x,y,t) - F(x-V_x, y-V_y, t-1)| \quad (1)$$

Под опорным подразумевается кадр, относительно которого осуществляется поиск векторов движения макроблоков текущего кадра. Иногда в качестве суммы норм принимают величину

$$SAD = \sum_{x,y=1,16} [F(x,y,t) - F(x-V_x, y-V_y, t-1)]^2 \quad (1.1)$$

Здесь F - значение яркости, (x, y) - пространственные координаты точки в кадре, t - временной индекс кадра, суммирование производится по всем точкам макроблока. Значение $\vec{V}$, для которого норма SAD имеет наименьшее значение, принимается за искомый вектор. Вектора движения ищутся методом полного перебора в некоторой ограниченной окрестности: $min < V_x, V_y < max$. Предполагая, что размер этой окрестности равен $\pm N$ по координатам x и y, получим для числа операций, необходимых для определения вектора движения одного макроблока, величину порядка $3 \cdot 256 (2N+1)^2$. На один пиксел макроблока количество операций составляет $3 \cdot (2N+1)^2$, что уже при N = 15 (значения векторов движения в пределах ± 15 точек) составляет значительную величину более 10^3 операций / пиксел.

Этот способ обычно используется в качестве эталонного для оценки качества других способов анализа векторов движения деталей в динамических изображениях.

Недостатком этого способа является требуемое большое количество вычислительных операций и в связи с этим низкое быстродействие.

Для уменьшения количества операций были предложены различные способы и методы, основанные на ограничении возможного числа кандидатов на вектора движения, в частности логарифмический и иерархический методы поиска векторов движения [1, с.92-96]. При логарифмическом поиске вектора движения поиск осуществляется в небольшой окрестности с шагом изменения значения вектора движения в 2 пиксела до тех пор, пока не будет найден локальный минимум SAD, после чего значение вектора движения уточняется с шагом в один пиксел в окрестности этого минимума. В 3-х шаговом способе поиск ведется первоначально по всей окрестности с шагом в 4 пиксела, затем вокруг найденного минимума с шагом в 2 пиксела, и, наконец, с шагом в один пиксел. В иерархическом способе первоначально проводится поиск векторов движения для специальным образом передискретизированного изображения, обладающего более грубым разрешением, потом значения найденных векторов движения уточняются при переходе к следующему уровню более высокого разрешения.

Недостатками этих ускоренных способов анализа являются существенные ошибки анализа движений деталей в динамических изображениях вследствие достаточно сложной зависимости SAD от V, которая практически всегда не является гладкой функцией. Функция $SAD(\vec{V})$ может иметь несколько экстремумов, что усложняет применение и значительно снижает эффективность данных методов.

Наиболее близким по технической сущности к заявляемому техническому решению является наиболее простой и точный способ [1] анализа векторов движения деталей в динамических изображениях, включающий преобразование последовательности кадров изображений в цифровую форму, запоминание дискретных отсчетов яркости текущего и соседнего по времени кадров, разбиение текущего кадра на макроблоки и поиск вектора движения каждого из макроблоков текущего кадра относительно опорного кадра посредством минимизации по рассматриваемому множеству векторов движения контрольной суммы данного макроблока, являющейся суммой норм попиксельной разности уровней в текущем и опорном кадрах.

Как уже было сказано, недостатком этого способа является требуемое большое количество вычислительных операций и в связи с этим низкое быстродействие. Этот недостаток не позволяет реально обеспечить большие зоны анализа движения, что приводит к снижению качества воспроизведения быстродвижущихся деталей в динамических изображениях.

В основу изобретения положена задача за счет сокращения объема вычислительных операций повысить быстродействие кодирующих устройств и производить анализ движения деталей динамических изображений в больших областях, сократив при этом объем сжатой информации и повысив качество воспроизведения быстродвижущихся деталей.

Эта задача решена посредством способа анализа векторов движения деталей в динамических изображениях, включающего преобразование последовательности кадров изображений в цифровую форму, запоминание дискретных отсчетов яркости текущего и опорного кадров, разбиение текущего кадра на макроблоки и поиск вектора движения каждого из макроблоков текущего кадра относительно опорного кадра посредством минимизации по рассматриваемому множеству векторов движения контрольной суммы данного макроблока, являющейся суммой норм попиксельной разности уровней в текущем и опорных кадрах. Согласно изобретению, все

RU 2137194 C1

RU 2137194 C1

множество пикселов рассматриваемого макроблока разбивают на участки, в каждом из которых определенным образом выбирают только один пиксел, и упомянутую контрольную сумму вычисляют лишь по выбранным упомянутым пикселам, при этом упомянутые пикселы в каждом из участков выбирают таким образом, что их уровни в смежных участках наиболее отличаются друг от друга.

Было бы целесообразно при использовании описанного выше способа анализа векторов движения деталей в динамических изображениях дополнительно для каждого текущего рассматриваемого значения вектора движения вычисление контрольной суммы по выбранным упомянутым пикселам производить в порядке убывания отклонения значения сигнала в каждом пикселе от среднего значения по всей совокупности упомянутых пикселов и прерыванием дальнейшего вычисления контрольной суммы в случае, когда она превышает значение минимальной контрольной суммы, найденной среди всех уже рассмотренных векторов движения.

На фиг. 1 изображен один из возможных вариантов устройства для реализации способа анализа векторов движения деталей в динамических изображениях; на фиг. 2 - 8 - диаграммы, поясняющие его работу.

Устройство, изображенное на фиг. 1, содержит подключенные к входу 1 параллельно блок синхронизации 2 и последовательно соединенные аналого-цифровой преобразователь 3, схему выделения 4 сигнала яркости, последовательно соединенные блок памяти 5 текущего кадра и блок памяти 6 опорного кадра, подключенные к выходам блока памяти текущего кадра последовательно соединенные блок памяти 7 текущего макроблока, блок преобразования 8 в порядковую статистику уровней пикселов макроблока и вычислитель 9 положений и уровней характерных пикселов макроблока, первые выходы которого соединены с блоком памяти 10 уровней выбранных пикселов, а вторые входы - через сумматор 11 к входам управления блока памяти опорного кадра 6, выходы которого через блок памяти 12 уровней пикселов сравнения подключены к первым входам блока 13 попиксельного вычитания уровней характерных пикселов текущего и опорного кадров, вторые входы которого подключены к вторым выходам блока памяти 10, а выходы - к блоку суммирования 14 модулей, выходы которого параллельно подключены к блоку сравнения сумм 15 и блоку вычисления 16 минимальной суммы и текущих векторов движения, вторые входы блока сравнения сумм 15 подключены к выходам блока 16, а выходы - непосредственно ко входам управления блока 14 и через счетчик перемещений 17 к второму входу сумматора 11, вторые выходы блока 16 подключены через блок памяти 18 векторов движения к выходам устройства 19, вторые выходы блока 18 соединены параллельно с дополнительными входами блока 16 и счетчика 17, выходы блока синхронизации 2 соединены с входами управления блоков 3 и 4, а также через блок управления 20 - с входами синхронизации работы блоков 5-18 устройства.

Сущность предполагаемого способа состоит в следующем.

Цель проводимых над сигналами динамического изображения преобразований состоит в резком сокращении объема вычислительных операций при увеличении зоны поиска векторов движения деталей в динамических изображениях, что позволяет существенно повысить быстродействие кодирующих устройств и производить анализ движения деталей динамических изображений в больших областях, сократив при этом объем сжатой информации и повысив качество воспроизведения быстродвижущихся деталей.

Изложение предлагаемого способа иллюстрируется на примере анализа вектора движения одного из макроблоков динамического изображения, приведенного на фиг. 2: фиг. 2а - текущий кадр, фиг. 2б - опорный кадр. На фиг. 3а приведена в увеличенном масштабе часть изображения текущего кадра в увеличенном масштабе с обведенным макроблоком, вектор движения которого определяется в зоне опорного кадра, изображенной на фиг. 3б. Макроблок имеет размер 16•16 пикселов, зона поиска движения - 64•64 пиксела. На фиг. 3б обозначено положение смещенного макроблока в опорном кадре.

В случае использования стандартного метода поиска вектора движения в указанной зоне поиска требуется использование вычислительной мощности в $3 \cdot 49^2 = 7203$ операций на каждый пиксел макроблока. Рельеф сигнала яркости макроблока изображен на фиг. 4а, а на фиг. 4б приведены цифровые значения пикселов этого макроблока.

В соответствии с предлагаемым способом макроблок разбивается на несколько участков, например на девять, как показано на фиг. 4б. В каждом из этих участков выбирается лишь один пиксел.

Один из вариантов выбора может быть связан с использованием порядковых статистик. В этом случае все значения пикселов $F_1 \ldots F_{256}$ располагаются в порядке возрастания уровней, как показано в таблице на фиг. 5. В этой таблице наряду со значениями уровней пикселов $F_{(1)} \le F_{(2)} \le \ldots\ldots F_{(255)} \le F_{(246)}$ приведена принадлежность их к соответствующим выбранным девяти участкам.

Для того, чтобы на точность оценки векторов движения в меньшей степени влияло наличие помех в сигнале яркости, в качестве первых двух пикселов принимаются пикселы, уровни в которых соответствуют следующему за минимальным и предшествующий максимальному (в данном случае - пиксел с уровнем $F_{(2)}$= 24, расположенный на 5-м участке макроблока и пиксел с уровнем $F_{(255)}$=236, расположенный на 7-м участке макроблока). Если эти пикселы оказываются расположенными в одном участке, то положение одного, например $F_{(2)}$, - фиксируется, и пикселы $F_{(254)}$, $F_{(253)}$ ... перебираются, пока участок окажется не тем, в котором располагается пиксел с уровнем $F_{(2)}$.

Пикселы в других участках выбираются отступая от $F_{(2)}$ в сторону увеличения индекса

RU 2137194 C1

RU 2137194 C1

и от $F_{(255)}$ в сторону уменьшения индекса, например, на равное количество пикселов $\frac{256}{9} \approx 28$. Следовательно, выбранными пикселами должны быть $F_{(30)}$=87, расположенный на 4-м участке и $F_{(227)}$=178, расположенный на 6-м участке.

Аналогично следующими пикселами выбираются $F_{(59)}$=112, расположенный на 9-ом участке, и $F_{(198)}$=164, расположенный на 1-м участке.

Далее выбираются пиксел с уровнем $F_{(88)}$=124, расположенный на 8-м участке, и пиксел с уровнем $F_{(169)}$=151, расположенный на 5-м участке. Учитывая, что на 5-м участке пиксел уже был выбран ранее ($F_{(2)}$), выбор осуществляется из двух пикселов $F_{(170)}$ и $F_{(168)}$ в данном случае с равными уровнями. Из них выбирается пиксел $F_{(170)}$, расположенный на 3-м участке.

И, наконец, в качестве претендента на выбранное значение пиксела можно указать $F_{(117)}$= 132. Но этот пиксел оказался также расположенным на 6-м участке и по этой причине подбирается пиксел с тем же уровнем $F_{(118)}$, расположенный на 2-м участке.

Таким образом,

на 1-м участке выбран пиксел с уровнем 164;

на 2-м участке выбран пиксел с уровнем 132;

на 3-м участке выбран пиксел с уровнем 151;

на 4-м участке выбран пиксел с уровнем 87;

на 5-м участке выбран пиксел с уровнем 24;

на 6-м участке выбран пиксел с уровнем 178;

на 7-м участке выбран пиксел с уровнем 236;

на 8-м участке выбран пиксел с уровнем 124;

на 9-м участке выбран пиксел с уровнем 112.

Для реализации этого метода подбора пикселов требуется использование примерно $\log_2 256 = 8$ вычислительных операций на каждый пиксел макроблока.

Возможен более простой метод выбора характерных пикселов в каждом из участков макроблока, например, путем чередования максимальных и минимальных значений уровней в смежных участках. Предположим, на заштрихованных на фиг. 4б участках выбираются пикселы с минимальными значениями уровней, а на остальных - с максимальными значениями уровней. Тогда

на 1-м участке будет выбран пиксел с уровнем 81;

на 2-м участке будет выбран пиксел с уровнем 196;

на 3-м участке будет выбран пиксел с уровнем 84;

на 4-м участке будет выбран пиксел с уровнем 221;

на 5-м участке будет выбран пиксел с уровнем 24;

на 6-м участке будет выбран пиксел с уровнем 237;

на 7-м участке будет выбран пиксел с уровнем 61;

на 8-м участке будет выбран пиксел с уровнем 188;

на 9-м участке будет выбран пиксел с уровнем 94.

Этот более простой метод подбора пикселов требует использования лишь одной вычислительной операции на пиксел.

Для поиска вектора движения $\vec{V} = (V_x, V_y)$ в предлагаемом способе рассматривается норма разницы сигналов яркости выбранных точек двух макроблоков в текущем и опорном кадрах SAD1 со сдвигом на вектор движения:

$$SAD1 = \sum_{\substack{(x,y)-\text{координаты выбранных точек в соответствующих участках макроблока}} |F(x,y,t)-F(x-V_x,y-V_y,t-1)| \quad (2)$$

Очевидно, что скорость анализа движения при использовании предлагаемого способа увеличивается в N/M раз, где N - число пикселов в макроблоке, M - количество участков, на которое поделен макроблок. В данном случае скорость анализа увеличивается в 256/9=28,44 раз.

Дополнительно скорость анализа движения макроблоков увеличивается в соответствии с п. 2 настоящего изобретения за счет определенного подбора вычисления контрольной суммы SAD 1.

В этом случае вычисляется среднее значение яркости выбранных пикселов макроблока

$$F_{ср} = \frac{1}{M} \sum_{\substack{(x,y)-\text{координаты выбранных точек в соответствующих участках макроблока}} F(x,y,t) \quad (3)$$

Затем вычисляются модули разностей $X(x,y) = |F(x,y,t)-F_{ср}|$,

где (x, y) - координаты выбранных точек в соответствующих участках макроблока, и контрольная сумма (2) вычисляется последовательно по координатам (x, y), для которых величины X(x, y) располагаются в убывающей последовательности.

В случае, если текущий выбранный вектор $\vec{V} = (V_x, V_y)$ не соответствует истинному вектору движения, для которого контрольная сумма минимальна, рассчитываемая контрольная сумма статистически достаточно быстро возрастает и в этом случае обычно не требуется производить вычисления по всем выбранным точкам макроблока. При этом в среднем по кадру скорость анализа движения макроблоков дополнительно возрастает в два раза.

Эффективность предлагаемого способа анализа движения иллюстрируется на фиг. 6-8.

На фиг. 6 приведены изображения разности текущего и опорного кадров после компенсации движения (а) и векторов движения (б), рассчитанных стандартным методом с использованием всех 256 пикселов макроблока.

На фиг. 7 приведены изображения разности текущего и опорного кадров после компенсации движения (а) и векторов

RU 2137194 C1

RU 2137194 C1

движения (б), рассчитанных предлагаемым способом с использованием лишь 9 пикселов макроблока.

На фиг. 8 приведены значения результирующей контрольной суммы до и после проведения анализа векторов движения; цифрой 1 обозначена кривая, характеризующая контрольную сумму до проведения анализа движения (без компенсации движения), цифрой 2 - кривая, характеризующая контрольную сумму при стандартном медленном анализе движения макроблоков с использованием 256 пикселов; цифрой 3 - кривая, характеризующая контрольную сумму при предложенном быстром способе анализа движения макроблоков с использованием лишь 9 пикселов.

Из визуального сравнения диаграмм фиг.6б и 7б видно, что анализ движения макроблоков с использованием стандартного и предлагаемого способов практически во всех областях кадра, насыщенных резко изменяющимся рельефом, осуществляется одинаково. Отличия возникают лишь в областях с малым изменением рельефа. Это подтверждается фиг. 8, на которой видно, что различие контрольных сумм оказывается малым (менее 10%).

Испытания различных динамических изображений показывают, что при увеличении числа выбранных участков, скажем, с 9 до 16 различие контрольных сумм с использованием стандартного и предлагаемого способа практически отсутствует.

Устройство для реализации предложенного способа работает следующим образом.

Предположим, на вход 1 устройства поступает аналоговый сигнал изображения, например полный цветовой телевизионный сигнал стандартных систем SECAM, PAL или NTSC. Со входа этот сигнал подается параллельно на блок синхронизации 2, выделяющий соответствующие сигналы и обеспечивающий формирование импульсов дискретизации, и на аналого-цифровой преобразователь АЦП 2, в котором дискретные отсчеты сигнала преобразуются в цифровой код, поступающий на схему выделения сигнала яркости 4, устраняющую цветовые поднесущие из полного цветового телевизионного сигнала. Выделение сигнала яркости необходимо, поскольку в соответствии со стандартами MPEG анализ движения деталей изображения осуществляется только при использовании яркостной компоненты изображения. Синхронизация работы схемы 4 осуществляется также дискретизирующими импульсами, поступающими с блока синхронизации 2.

Цифровой поток сигнала яркости со схемы 4 последовательно поступает на блок памяти 5 текущего кадра и блок памяти 6 опорного (предыдущего) кадра, в которых запоминаются дискретные отсчеты сигналов яркости текущего кадра, анализ движения макроблоков которого осуществляется относительно соответствующих структур опорного кадра.

Выходы блока 5 соединены со входами блока памяти макроблока 7, анализ движения которого производится. В этом блоке запоминаются значения рельефа сигнала яркости макроблока 16 •16 пикселов. После вычисления вектора движения 1-го макроблока, соответствующего левому верхнему углу изображения, в память этого блока вводится рельеф сигнала яркости следующего за ним макроблока. Отсчет номера макроблока осуществляется обычно слева направо и сверху вниз.

После запоминания уровней яркости всех пикселов макроблока они в блоке 8 преобразуются в порядковую статистику.

Обычно такое преобразование производится с использованием быстрого алгоритма, требующего для N чисел $N\log_2 N$ операций. При построении порядковой статистики одновременно фиксируются координаты пикселов, которым соответствует каждое их значение.

По значениям порядковой статистики в соответствии с упомянутым выше способом с помощью вычислителя положений и уровней 9 выбираются расположенные в выбранных участках по одному пикселу, по значениям, например, номеров порядковых статистик, приблизительно равноудаленных друг от друга. Принцип расчета поясняется для выбранного макроблока фиг.4а с помощью таблиц на фиг. 4б и 5. Рассчитанные выбранные пикселы соответствуют координатам, приведенным в табл. 1.

В случае использования описанного выше второго варианта выбора точек их координаты будут иными (см. табл. 2)

Таким образом, в вычислителе 9 выделяются значения уровней яркости выбранных пикселов и их координаты x и y.

В этом же блоке осуществляется преобразование выбранных пикселов в последовательность по мере убывания отклонений значений уровней пикселов от их среднего значения, для этой цели в первом варианте вычисляется среднее значение уровней

$F_{ср}$ =(164+132+151+87+24+178+236+124+112)/9 ≈ 134, а затем модули

$X(i) = X(x,y) = |F(x,y,t)-F_{ср}|$:

X(1) = 30; X(2) = 2; X(3) = 17; X(4)= 47; X(5)= 110;

X(6) = 44; X(7) = 102; X(8) = 10; X(9) = 22.

При этом выбранные пикселы представляются и запоминаются в вычислителе 9 в последовательности уменьшения модулей разностей (см. табл. 3).

При выборе пикселов по второму варианту среднее значение яркости будет равно:

$F_{ср}$ =(81+196+84+221+24+237+61+188+94)/9≈132.

Модули разностей при этом равны:

X(1)=51; X(2)=64; X(3)=48; X(4)=91; X(5)=108;

X(6) = 105; X(7) = 71; X(8) = 56; X(9) = 38, и выбранные пикселы представляются и запоминаются в вычислителе 9 в последовательности (см. табл. 4).

Далее уровни сигналов яркости в рассчитанной пиксельной последовательности запоминаются в блоке памяти 10, а их координаты (x, y) - через сумматор 11 поступают на вход управления блока памяти опорного кадра 6, обеспечивая запись в блоке 12 памяти уровней пикселов сравнения уровней пикселов с координатами

RU 2137194 C1

RU 2137194 C1

$(x\text{-}V_{OX}\text{-}V_{x,y}\text{-}V_{OY}\text{-}V_Y)$,

где (V_{OX}, V_{OY})- координаты начального вектора смещения, которые определяются, например, из результатов оценки движения соответствующих макроблоков в предыдущем кадре или из иных соображений,

(V_X, V_Y) - координаты текущего вектора смещения макроблока, при изменении которых производится анализ контрольной суммы разностей модулей уровней выбранных пикселов в макроблоке и пикселов в опорном кадре.

С блоков памяти 10 и 12 в указанной выше последовательности значения соответствующих уровней пикселов поступают на блок попиксельного высчитания 13 и с его выхода - на блок суммирования модулей 14. Блоки 13 и 14 совместно реализуют операцию, определяемую соотношением (2).

Последовательно вычисляемая в блоке 14 сумма поступает параллельно на блок сравнения сумм 15 и блок вычисления 16 минимальной суммы и соответствующих ей векторов движения. Первоначально при переходе от одного макроблока к другому в блоке 16 запоминается рассчитанный ранее начальный вектор движения (V_{OX}, V_{OY}), а сумма устанавливается нереально большой величиной.

При вычислении суммы для начального вектора движения эта сумма запоминается в блоке 16 и становится опорной для последующих расчетов. Если эта сумма при первом же расчете оказывается ниже некоторой установленной малой величины, то считается, что вектор движения найден. В противном случае по окончании расчета суммы при нулевом векторе (V_X, V_Y) блок сравнения 15 выдает команду на изменение координат вектора движения в счетчик перемещений 17 и обнуление результатов суммирования в блок 14. Обычно для обеспечения эффективного анализа вектора движения перемещение текущего вектора движения осуществляется от точки (V_{OX}, V_{OY}) по расходящейся спирали в пределах заданной зоны анализа.

В счетчике 17 формируется вектор $(-V_{OX} - V_X, - V_{OY} - V_Y)$, который затем суммируется в схеме 11 с текущими значениями координат выбранных характерных пикселов макроблока. Полученные координаты определяют запись уровней яркости соответствующих пикселов из блока 6 в блок 12, и процесс анализа вектора движения продолжается до тех пор, пока не найден будет истинный вектор движения макроблока (когда минимальная сумма, вычисляемая блоком 16, не станет меньше некоторой установленной минимальной величины), либо пока не будет проанализирована вся заданная зона анализа. Последний случай означает, что подобного макроблока в опорном кадре не существует.

Ускорение процесса анализа движения в соответствии с п. 2 настоящего изобретения определяется непрерывным сравнением текущей суммы и минимальной суммы в блоке 15. Если текущая сумма (до окончания процесса суммирования по всем выбранным пикселам) превышает ранее найденную минимальную сумму, то процесс суммирования прекращается и счетчик 17 перемещений изменяет координаты текущего вектора движения макроблока.

Результаты вычисления векторов движения, определяемые в блоке 16, запоминаются блоком памяти 18 векторов движения из которого результаты расчетов векторов движения, например в предыдущих кадрах (V_{OX}, V_{OY}), используются блоками 16 и 17 в качестве исходных векторов. С блока 18 координаты векторов движения подаются на цифровые выхода устройства 19.

Выполнение операций различных блоков устройства в описанной выше последовательности задается блоком управления, синхронизируемым импульсами с выхода блока синхронизации.

Источники информации

1. K. R.Rao, J.J. Hwang, "Techniques and Standards for Image, Video and Audio Coding", 1996, Prentice-Hall PTR, ISBN 0-13-309907-5.

Формула изобретения:

1. Способ анализа векторов движения деталей в динамических изображениях, включающий преобразование последовательности кадров изображений в цифровую форму, запоминание дискретных отсчетов яркости текущего и опорного кадров, разбиение текущего кадра на макроблоки и поиск вектора движения каждого из макроблоков текущего кадра относительно опорного кадра посредством минимизации по рассматриваемому множеству векторов движения контрольной суммы данного макроблока, являющейся суммой норм попиксельной разности уровней в текущем и опорных кадрах, отличающийся тем, что все множество пикселов рассматриваемого макроблока разбивают на участки, в каждом из которых определенным образом выбирают только один пиксел, и упомянутую контрольную сумму вычисляют лишь по выбранным упомянутым пикселам, при этом упомянутые пикселы в каждом из участков выбирают таким образом, что их уровни в смежных участках наиболее отличаются друг от друга.

2. Способ анализа векторов движения по п. 1, отличающийся тем, что для каждого текущего рассматриваемого значения вектора движения вычисление контрольной суммы по выбранным упомянутым пикселам производят в порядке убывания отклонения значения сигнала в каждом пикселе от среднего значения по всей совокупности упомянутых пикселов и прерывают дальнейшее вычисление контрольной суммы в случае, когда она превышает значение минимальной контрольной суммы, найденной среди всех уже рассмотренных векторов движения.

RU 2137194 C1

RU 2137194 C1

Таблица 1

№ участка	Координаты пиксела		Уровень сигнала яркости
	x	*y*	
1	5	5	164
2	8	1	132
3	15	1	151
4	5	10	87
5	6	11	24
6	15	8	178
7	2	15	236
8	6	13	124
9	13	16	112

Таблица 2

№ участка	Координаты пиксела		Уровень сигнала яркости
	x	*y*	
1	2	1	81
2	11	1	196
3	14	3	84
4	1	9	221
5	6	11	24
6	12	9	237
7	3	12	61
8	8	13	188
9	16	16	94

Таблица 3

№ участка	Координаты пиксела		Уровень сигнала яркости	Модуль разности
	x	*y*		
5	6	11	24	110

RU 2137194 C1

RU 2137194 C1

Продолжение таблицы 3

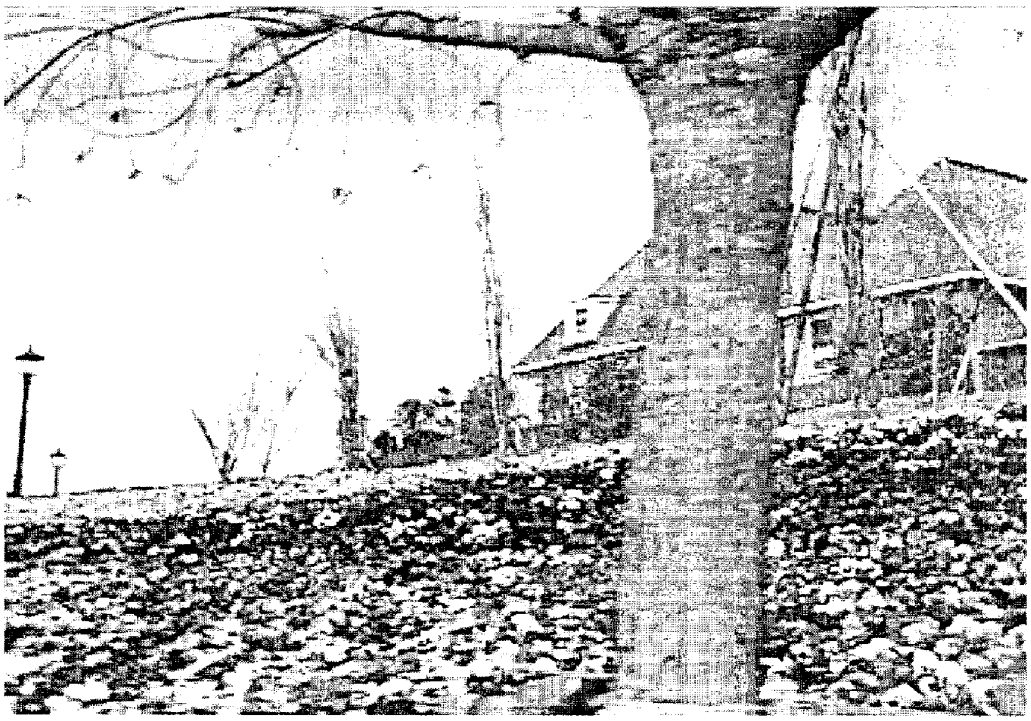
7	2	15	236	102
4	5	10	87	47
6	15	8	178	44
1	5	5	164	30
9	13	16	112	22
3	15	1	151	17
8	6	13	124	10
2	8	1	132	2

Таблица 4

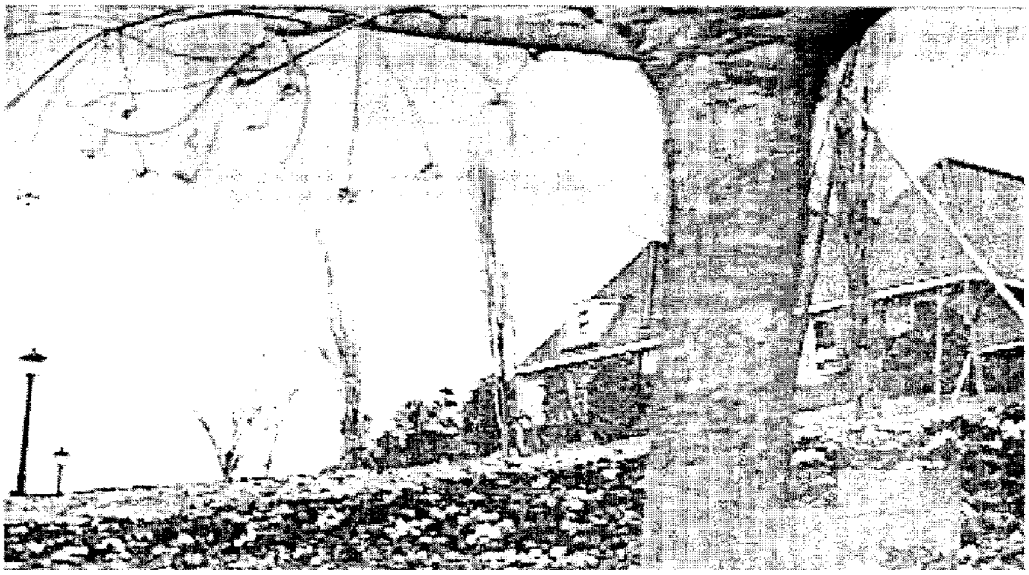
№ участка	Координаты пиксела		Уровень сигнала яркости	Модуль разности
	x	*y*		
5	6	11	24	108
6	12	9	237	105
4	1	9	221	91
7	3	12	61	71
2	11	1	196	64
8	8	13	188	56
1	2	1	81	51
3	14	3	84	48
9	16	16	94	38

RU 2137194 C1

RU 2137194 C1

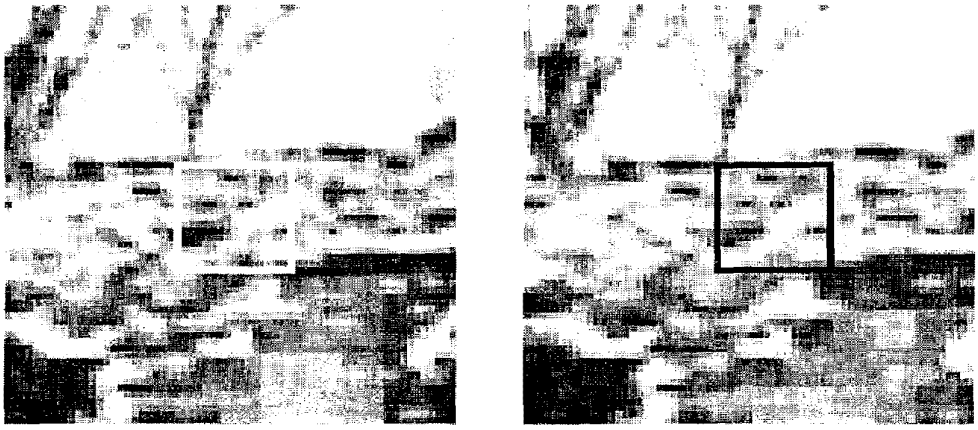


(a)



RU 2137194 C1

RU 2137194 C1

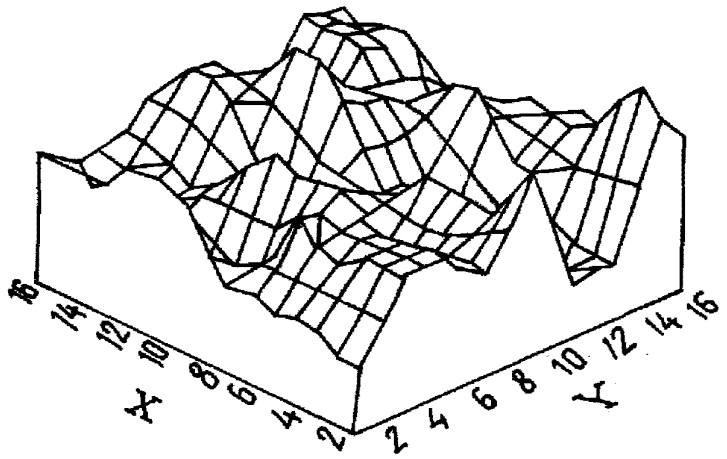


(а) (б)

Фиг.3.

RU 2137194 C1

RU 2137194 C1



а

Таблица уровней пикселов в макроблоке, разделенном на 9 участков

y \ x	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
	Участок № 1					Участок № 2						Участок № 3				
1	82	81	97	102	94	105	100	132	171	189	196	181	151	140	151	151
2	126	125	127	125	127	127	112	104	117	137	150	144	125	117	125	135
3	167	165	168	142	153	174	115	62	64	87	118	121	86	84	106	117
4	166	163	158	154	161	165	147	126	123	127	125	114	101	103	115	128
5	166	160	156	156	164	183	178	174	186	168	138	115	100	119	131	131
	Участок № 4					Участок № 5						Участок № 6				
6	140	140	135	128	125	132	145	153	147	128	117	126	143	157	159	157
7	120	122	116	104	75	80	117	136	123	80	82	139	186	203	195	170
8	164	149	131	120	115	115	120	126	134	143	160	182	191	187	178	170
9	221	179	136	142	147	151	125	111	152	197	235	237	204	172	166	163
10	142	120	100	90	87	85	88	109	148	182	195	188	170	150	132	125
11	70	61	55	39	23	24	41	91	151	180	171	141	139	126	95	78
	Участок № 7					Участок № 8						Участок № 9				
12	78	63	61	64	69	80	103	136	159	163	153	148	153	151	139	131
13	68	61	68	85	120	124	161	188	179	157	134	143	173	186	177	180
14	146	143	138	134	140	148	158	162	151	133	124	140	166	174	170	166
15	210	236	210	169	172	161	163	148	124	110	97	127	174	169	155	161
16	181	189	176	145	120	113	121	124	112	94	86	97	112	110	99	94

(б)

Фиг. 4.

RU 2137194 C1

RU 2137194 C1

Уровень пиксела	23	[24]	39	41	55	61	61	61	62	63	64	64	68	68	69	70
Номер участка	4	[5]	4	5	4	7	7	4	2	7	2	7	7	7	7	4

75	78	78	80	80	80	81	82	82	84	85	85	86	86	[87]	87
4	7	6	5	5	8	1	1	5	3	7	5	3	8	[4]	2

88	90	91	94	94	94	95	97	97	97	99	100	100	100	101	102
5	4	5	1	8	9	6	1	9	8	9	3	4	2	3	1

103	103	104	104	105	106	109	110	110	111	112	[112]	112	113	114	115
3	8	4	2	2	3	5	8	9	5	8	[9]	2	8	3	4

115	115	115	115	116	117	117	117	117	117	118	119	120	120	120	120
2	5	3	3	4	3	5	3	5	2	2	3	4	4	7	5

120	120	121	121	122	123	123	124	[124]	124	124	125	125	125	125	125
7	4	3	8	4	2	5	8	[8]	8	8	4	6	5	1	3

125	125	125	126	126	126	126	126	127	127	127	127	127	128	128	128
1	3	2	1	6	2	5	6	9	2	1	2	1	3	4	5

131	131	131	131	132	{132}	[132]	133	134	134	134	135	135	136	136	136
3	4	9	3	5	{6}	[2]	8	7	5	8	3	4	8	4	5

137	138	138	139	139	139	140	140	140	140	140	141	142	142	142	143
2	2	7	6	6	9	3	4	4	7	9	6	4	4	1	6

143	143	143	144	145	145	146	147	147	147	148	148	148	148	149	150
7	5	9	3	5	7	7	5	2	4	5	8	9	8	4	6

150	151	151	151	151	151	{151}	[151]	152	153	153	153	153	154	155	156
2	3	3	8	5	9	{5}	[3]	5	5	1	8	9	1	9	1

156	157	157	157	158	158	159	159	160	160	161	161	161	161	162	163
1	8	6	6	1	8	8	6	5	1	8	8	1	9	8	1

163	163	163	[164]	164	165	165	166	166	166	166	166	167	168	168	169
8	6	8	[1]	4	2	1	6	9	9	1	1	1	1	2	9

169	170	170	170	170	171	171	172	172	174	174	174	174	176	177	178
7	6	6	9	6	2	5	7	6	9	2	2	9	7	9	9

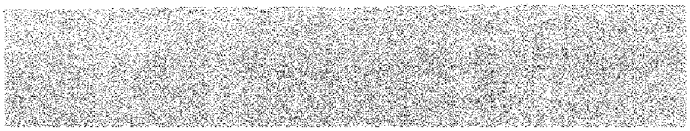
178	[178]	179	179	180	180	181	181	182	182	183	186	186	186	187	188
2	[6]	4	8	9	5	3	7	6	5	2	6	2	9	6	6

188	189	189	191	195	195	196	197	203	204	210	210	221	235	[236]	237
8	7	2	6	5	6	2	5	6	6	7	7	4	5	[7]	6

Фиг. 5.

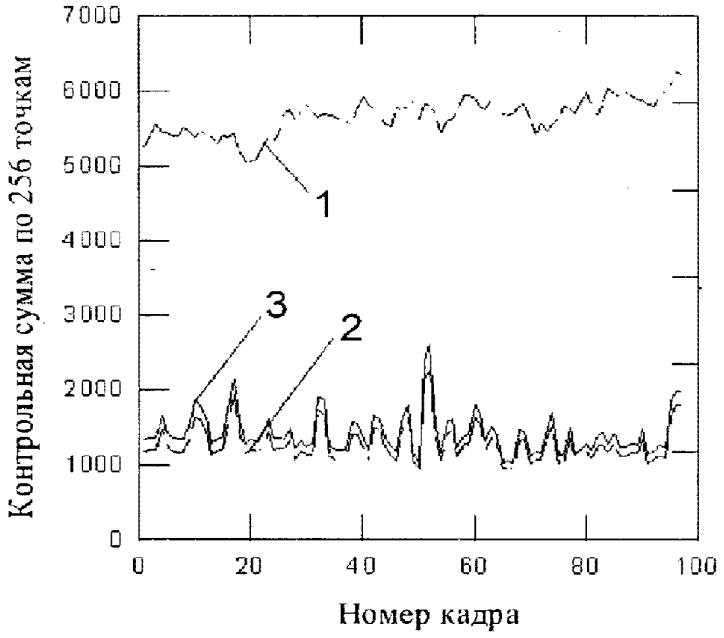
RU 2137194 C1

RU 2137194 C1



RU 2137194 C1

RU 2137194 C1



Фиг. 8.

RU 2137194 C1

Electronic Patent Application Fee Transmittal

Application Number:	11338934
Filing Date:	25-Jan-2006
Title of Invention:	Grouping of image frames in video coding
First Named Inventor/Applicant Name:	Miska Hannuksela
Filer:	Steven R. Funk/Tracey Dotter
Attorney Docket Number:	TPO.004.C2

Filed as Large Entity

Utility under 35 USC 111(a) Filing Fees

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Basic Filing:				
Pages:				
Claims:				
Miscellaneous-Filing:				
Petition:				
Patent-Appeals-and-Interference:				
Post-Allowance-and-Post-Issuance:				
Extension-of-Time:				

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Miscellaneous:				
Submission- Information Disclosure Stmt	1806	1	180	180
		Total in USD ($)		**180**

Electronic Acknowledgement Receipt

EFS ID:	8190380
Application Number:	11338934
International Application Number:	
Confirmation Number:	6610
Title of Invention:	Grouping of image frames in video coding
First Named Inventor/Applicant Name:	Miska Hannuksela
Customer Number:	76385
Filer:	Steven R. Funk
Filer Authorized By:	
Attorney Docket Number:	TPO.004.C2
Receipt Date:	10-AUG-2010
Filing Date:	25-JAN-2006
Time Stamp:	17:33:08
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	yes
Payment Type	Deposit Account
Payment was successfully received in RAM	$180
RAM confirmation Number	3997
Deposit Account	503581
Authorized User	
The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows: Charge any Additional Fees required under 37 C.F.R. Section 1.17 (Patent application and reexamination processing fees)	

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Transmittal Letter	IDS-10-TPO004C2.pdf	55426 3a289480cadfd3eef89a80ef685a2fe46b4ae296	no	2
Warnings:					
Information:					
2	Information Disclosure Statement (IDS) Filed (SB/08)	1449-10-TPO004C2.pdf	53435 5ecde266655871ab0f61200da8cf90c9f89cb2a3	no	1
Warnings:					
Information:					
This is not an USPTO supplied IDS fillable form					
3	NPL Documents	FH-TAPA004PA.pdf	999816 e0f3b198e0372dfeee0d292f6afef40141aed93c	no	27
Warnings:					
Information:					
4	NPL Documents	Illgner.pdf	524122 72978d8c3d123a922894913ff66ff592b86674fa	no	16
Warnings:					
Information:					
5	NPL Documents	Informationtechnology.pdf	1641793 d76b603427c9443ca71ff18667af1f1834b69a97	no	174
Warnings:					
Information:					
6	NPL Documents	RUAllowance.pdf	4724947 318f5d65aaeeedf8f2b368397e91b944baced92a	no	11
Warnings:					
Information:					
7	Foreign Reference	RU2137194C1.pdf	973318 31ad8c486895acf06c549d3f83a00eff14e65d24	no	15
Warnings:					
Information:					
8	NPL Documents	TPO004C1-OA06172010.pdf	976305 3669dce1f618145853a25ebde8d2ffbab41c3937	no	30
Warnings:					
Information:					

9	Fee Worksheet (PTO-875)	fee-info.pdf	30215 fa100b9ba2d268b097e5f4bec88273347ac5abcb	no	2

Warnings:	
Information:	
Total Files Size (in bytes):	9979377

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111

If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371

If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office

If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

IN THE UNITED STATES PATENT & TRADEMARK OFFICE

Applicant:	HANNUKSELA	Examiner:	Lee, S.
Serial No.:	11/338,934	Group Art Unit:	2611
Filed:	January 25, 2006	Docket No.:	TPO.004.C2
Title:	GROUPING OF IMAGE FRAMES IN VIDEO CODING		

SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT
UNDER 37 C.F.R. §1.97(c)

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

Applicant respectfully submits the items of information on the enclosed Form 1449 for the attention of the Examiner in the above-identified application.

This statement should be considered because it is submitted after the mailing date of the first Office Action on the merits but before the mailing date of a final Office Action under 37 C.F.R. §1.113 or a Notice of Allowance under 37 C.F.R. §1.311. Please charge Deposit Account 50-3581 (TPO.004.C2) the amount of $180.00 under 37 C.F.R. §1.17(p) for consideration of the item.

In accordance with 37 C.F.R. §1.98(a)(2), only a copy of each foreign document or non-U.S. patent/application listed on the enclosed Form 1449 is provided.

Please note that any notations or markings on the attached documents do not reflect particular relevance, or lack thereof, to the present application, nor were they necessarily made by anyone affiliated with the prosecution of the present application.

No representation is made that the reference is "prior art" within the meaning of 35 U.S.C. §§102 and 103. In addition, Applicant does not represent that a reference has been thoroughly reviewed or that any relevance of any portion of a reference is intended, and reserves the right to establish otherwise under 37 C.F.R. §1.131 or others.

Consideration of the items listed is respectfully requested. According to MPEP § 609, Applicant requests that the Examiner return a copy of the attached Form 1449, marked as being considered and initialed by the Examiner, to the undersigned with the next official communication.

Respectfully submitted,

HOLLINGSWORTH & FUNK, LLC
8500 Normandale Lake Blvd., Suite 320
Minneapolis, MN 55437
952.854.2700

By: /Steven R. Funk/
Steven R. Funk
Reg. No. 37,830

IN THE UNITED STATES PATENT & TRADEMARK OFFICE

Applicant:	HANNUKSELA	Examiner:	Lee, S.
Serial No.:	11/338,934	Group Art Unit:	2611
Filed:	January 25, 2006	Docket No.:	TPO.004.C2
Confirmation No.:	6610	Customer No.:	76385
Title:	GROUPING OF IMAGE FRAMES IN VIDEO CODING		

CERTIFICATE UNDER 37 CFR 1.8: The undersigned hereby certifies that this paper is being electronically transmitted by EFS-WEB to the United States Patent and Trademark Office on September 17, 2010.

By: /Erin Nichols Matkaiti/
Erin Nichols Matkaiti

AMENDMENT & RESPONSE UNDER 37 C.F.R. § 1.111

Mail Stop Amendment
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

In response to the Office action mailed May 25, 2010, setting a three (3) month shortened statutory period for response, Applicant submits the present amendment and response under 37 C.F.R. § 1.111.

Claims 1, 2, 4-6, 8-11, 13, and 14 remain pending in the application. Claims 1, 5, 6, 8-10, 13, and 14 have been amended. No claims have been added, and claims 3, 7, 12, 15, and 16 have been canceled. Applicant respectfully requests reconsideration and allowance of all pending claims.

AMENDMENTS TO THE CLAIMS are reflected in the listing of claims which begins on page 2 of this paper.

REMARKS/ARGUMENTS begin on page 6.

AMENDMENTS TO THE CLAIMS

This listing of claims replaces all prior versions of claims in the application.

1. (Currently amended) A method for encoding a video sequence comprising an independent sequence of image frames, wherein at least one reference image frame is predictable from at least one previous image frame, and wherein the at least one reference image frame ~~that~~ is earlier than the at least one previous ~~reference~~ image frame in decoding order, the method comprising:

encoding into the video sequence an indication of at least one image frame, which is the first image frame, in decoding order, of the independent sequence, said at least one reference image frame being included in the sequence;

encoding identifier values for image frames according to a numbering scheme; and

resetting the identifier value for the indicated first image frame of an independent sequence.

2. (Original) A method according to claim 1, further comprising:

encoding the indication into the video sequence as a separate flag included in the header of a slice.

3. (Canceled)

4. (Original) A method according to claim 1, further comprising:

encoding into the video sequence an identifier value for the independent sequence.

5. (Currently amended) A video encoder for encoding a video sequence comprising an independent sequence of image frames, wherein at least one reference image frame is predictable from at least one previous image frame, and wherein the at least one reference frame ~~that~~ is earlier than the at least one previous ~~reference~~ image frame in decoding order, the video encoder being ~~arranged~~configured to:

encode into the video sequence an indication of at least one image frame, which is the first image frame, in decoding order, of the independent sequence, said at least one reference image frame being included in the sequence;

encode identifier values for image frames according to a numbering scheme; and

reset the identifier value for the indicated first image frame of an independent sequence.

6. (Currently amended) A video encoder according to claim 5, wherein the video encoder is ~~arranged~~configured to

encode the indication into the video sequence as a separate flag included in the header of a slice.

7. (Canceled)

8. (Currently amended) A video encoder according to claim 5, wherein the video encoder is ~~arranged~~configured to

encode into the video sequence an identifier value for the independent sequence.

9. (Currently amended) A computer program product, stored on a non-transitory computer readable medium and executable in a data processing device, for encoding a video sequence comprising an independent sequence of image frames, wherein at least one reference image frame is predictable from at least one previous image frame [[that]]and wherein the at least one reference frame is earlier than the at least one previous ~~reference~~ image frame in decoding order, the computer program product comprising:

a computer program code for encoding into the video sequence an indication of at least one image frame, which is the first image frame, in decoding order, of the independent sequence, said at least one reference image frame being included in the sequence;

a computer code for encoding identifier values for image frames according to a numbering scheme; and

a computer code for resetting the identifier value for the indicated first image frame of an independent sequence.

10. (Currently amended) A method for decoding a compressed video sequence, wherein at least one reference image frame is predictable from at least one previous image frame [[that]]and wherein the at least one reference frame is earlier than the at least one previous ~~reference~~ image frame in decoding order, the method comprising:

decoding from the video sequence an indication of at least one image frame, which is the first image frame, in decoding order, of an independent sequence, said sequence including the at least one reference image frame; [[and]]

starting the decoding of the video sequence from said first image frame of the independent sequence, whereby the video sequence is decoded without prediction from any image frame decoded prior to said first image frame;

decoding identifier values for image frames according to a numbering scheme; and

resetting the identifier value for the indicated first image frame of an independent sequence.

11. (Original) A method according to claim 10, wherein the indication is a separate flag included in the header of a slice.

12. (Canceled)

13. (Currently amended) A video decoder for decoding a compressed video sequence, wherein at least one reference image frame is predictable from at least one previous image frame, and wherein the at least one reference image frame ~~that~~ is earlier than the at least one previous ~~reference~~ image frame in decoding order, the video decoder being ~~arranged~~configured to:

decode from the video sequence an indication of at least one image frame, which is the first image frame, in decoding order, of an independent sequence, said sequence including the at least one reference image frame; [[and]]

start the decoding of the video sequence from said first image frame of the independent sequence, whereby the video sequence is decoded without prediction from any image frame decoded prior to said first image frame; and

decode identifier values for image frames according to a numbering scheme; and

reset the identifier value for the indicated first image frame of an independent sequence.

14. (Currently amended) A computer program product, stored on a non-transitory computer readable medium and executable in a data processing device, for decoding a compressed video sequence, wherein at least one reference image frame is predictable from at least one previous image frame, and wherein the at least one reference image frame ~~that~~ is earlier than the at least one previous ~~reference~~ image frame in decoding order, the computer program product comprising:

a computer program code for decoding from the video sequence an indication of at least one image frame, which is the first image frame, in decoding order, of an independent sequence, said sequence including the at least one reference image frame; [[and]]

a computer program code for starting the decoding of the video sequence from said first image frame of the independent sequence, whereby the video sequence is decoded without prediction from any image frame decoded prior to said first image frame;

a computer program code for decoding identifier values for image frames according to a numbering scheme; and

a computer program code for resetting the identifier value for the indicated first image frame of an independent sequence.

15-16. (Canceled)

REMARKS/ARGUMENTS

The rejections presented in the Office action dated May 25, 2010 (hereinafter Office Action), have been considered. Reconsideration of the pending claims and allowance of the application in view of the present response is respectfully requested.

Claims 3, 7, 12, 15, and 16 have been canceled, without prejudice, rendering the respective rejections of these claims moot. Applicant accordingly requests that the rejections of each of these claims be withdrawn.

With respect to the objection to claims 1, 2, 4, 5, 6, 8, 10, and 11, the term "the previous reference image frame" in independent claims 1, 5, and 10 has been amended to read "the previous image frame" in order to provide antecedent basis for the term. Applicant accordingly requests that the objection be removed.

With respect to the objection to Fig. 1, Applicant respectfully traverses. In contrast to the assertion that the figure includes "blank boxes", each of the boxes of Fig. 1 is labeled with a reference numeral in accordance with MPEP § 608.02. Moreover, no requirement has been cited for the request to label the boxes with text. Since Fig. 1 complies with the drawing requirements, and in an effort to maintain consistency with the parent application, Applicant requests that the objection be removed.

With respect to the § 101 rejection of claims 5-8 and 13, Applicant respectfully traverses. The Office action incorrectly asserts that the claimed encoder and decoder are "pure software". While the specification refers to embodiments involving computer readable media storing program code (*e.g.*, claims 9 and 14), the claimed encoder and decoder are devices constituting statutory subject matter. The terms "encoder" and "decoder" are well-known in the art to refer to a device or circuit. Notably, the specification illustrates structural components of the encoder (*e.g.*, compressor 108 and multiplexer 110) and decoder (*e.g.*, demultiplexer 116 and decompressor 118). Paragraph [0095] teaches examples of a decoder including a low bit rate video decoder as described in the ITU-T recommendation H.263 or H.26L. Moreover, paragraph [0096] describes positioning of the encoder and decoder that would not be possible if they were "pure

software". Contrary to the assertion, the claimed encoder and decoder are devices constituting statutory subject matter. Applicant accordingly requests that the rejection be withdrawn.

With respect to the § 101 rejection of claims 9 and 14, the claims have been amended to indicate that the claimed computer readable media are non-transitory in accordance with the Examiner's suggestion. Applicant notes that non-transitory computer readable media comprise all computer readable media, with the sole exception being a transitory, propagating signal. Since an interpretation of the claims is now limited to statutory subject matter, Applicant submits that the rejection has been overcome and accordingly requests that the rejection be withdrawn.

Applicant respectfully traverses the § 112, first paragraph, rejection of claims 5, 6, 8, and 13 because the claims are not "single means" claims. First, none of these claims uses a "means recitation" or "means plus function" language. While a claim dependent upon a recited property could result in a fact situation similar to that discussed in the cited *Hyatt* decision when the claim covers every conceivable structure for achieving the stated property, such is not the case for claims 5, 6, 8, and 13. In contrast the claims are limited to the structures of an encoder (claims 5, 6, and 8) and of a decoder (claim 13), which are well-accepted and understood terms in the art, for achieving the stated properties of the respective claims. Therefore, the claims are not unduly broad and are enabled by the specification. Applicant accordingly requests that the rejection be withdrawn.

Applicant respectfully traverses the § 102(e) rejection of claim 1 in view of U.S. Publication No. 2003/0063806 by Kim *et al*. (hereinafter "Kim") because the Office action acknowledges that Kim fails to teach or suggest each of the claimed limitations. In the § 103(a) rejection of claim 1 at page fifteen, the Office action states that "Kim fails to explicitly disclose an indication of the first I frame". Since Kim admittedly fails to correspond to each of the limitations of at least claim 1, the § 102(e) rejection is improper. Applicant accordingly requests that the rejection be withdrawn.

In an effort to facilitate prosecution and without acquiescing to characterizations of the asserted art, Applicant's claimed subject matter, or to the applications of the asserted art

or combinations thereof to Applicant's claimed subject matter, the independent claims have been amended to include the respective limitations of dependent claims 3, 7, and 12; therefore, these changes do not introduce new matter. Since the Office action acknowledges that Kim and U.S. Patent No. 6,094,456 to Ueda fail to teach or suggest each of the limitations of dependent claims 3, 7, and 12, each of the rejections based on these references is overcome. Each of the claims is further believed to be patentable over the asserted references for the reasons set forth below.

In rejecting dependent claims 3, 7, and 12, the Office action relies upon the teachings of U.S. Patent No. 4,390,966 to Kawashima *et al.* (hereinafter "Kawashima") as teaching the claimed resetting of the identifier value for the indicated first image frame of an independent sequence; however, this reliance is misplaced. Kawashima merely teaches that a sequence controller system is capable of setting and resetting flags according to predetermined user instructions. Kawashima's flags fail to at least correspond to the claimed identifier value since Kawashima is silent with respect to the flags being associated with an image frame of a coded video sequence. Also, Kawashima's flags are toggle flags for use in a multi CPU environment where each controller can run sequentially stored code in a parallel environment. There is no correspondence to the claimed resetting of an encoded identifier value for a first image frame of a video sequence. Without correspondence to each of the claimed limitations, the § 103(a) rejections of the subject matter of dependent claims 3, 7, and 12 is improper. Applicant accordingly requests that the rejections be withdrawn.

Further, the asserted modification of the combination of Kim and Ueda with the teachings of Kawashima is improper since the asserted modification is illogical and the teachings of Kawashima are not analogous to the field of Applicant's endeavor. First, the Office action asserts that it would be desirable to reset the identifier value since the video sequence is independent because it would simplify the identifier value and reset the identifier value. No explanation or evidence has been provided to indicate that resetting Kawashima's flags would simplify a value. Moreover, it is illogical to assert that it is

desirable to reset an identifier value because it would reset the identifier value. Thus, the proposed modification to the combination of Kim and Ueda is illogical and unsupported.

Second, the teachings of Kawashima are nonanalogous art as they are directed to very different technology than the field of video coding. In contrast, Kawashima is directed to the field of multi-core processing. Since Kawashima is in a field different than Applicant's endeavor, to be considered reasonably pertinent, Kawashima would need to address a problem addressed by the application at issue (MPEP § 2141.01(a)). However, the Office action fails to assert or identify any such relationship. Kawashima, directed to very different technology and failing to address video coding, would not have logically commended itself to an inventor of the instant application. Thus, Kawashima is nonanalogous art and fails to support the § 103(a) rejections. Applicant respectfully maintains the request that the § 103(a) rejections based upon Kawashima be withdrawn.

It should also be noted that Applicant does not acquiesce to the Examiner's statements or conclusions concerning what would have been obvious to one of ordinary skill in the art, obvious design choices, common knowledge at the time of Applicant's invention, officially noticed facts, and the like. Applicant reserves the right to address in detail the Examiner's characterizations, conclusions, and rejections in future prosecution.

With respect to the provisional rejections of claims 1-4 under the judicially-created doctrine of obviousness-type double patenting over co-pending and commonly-owned patent application no. 10/306,942 (hereinafter "the '942 application"), Applicant will consider filing a Terminal Disclaimer once the instant application is in condition for allowance or the '942 application issues as a patent. However, Applicant respectfully asserts that, in view of the amendments to the claims and arguments made above, the Examiner is compelled to withdraw the rejections of the claims. Once withdrawn, the only rejection remaining in the instant application would be the provisional obviousness-type double patenting rejection. In view of MPEP § 804 I(B), Applicant respectfully requests that the provisional obviousness-type double patenting rejection be withdrawn and that the instant application be permitted to issue as a patent.

Authorization is given to charge Deposit Account No. 50-3581 (TPO.004.C2) any necessary fees for this filing. If the Examiner believes it necessary or helpful, the undersigned attorney of record invites the Examiner to contact the undersigned attorney to discuss any issues related to this case.

Respectfully submitted,

HOLLINGSWORTH & FUNK, LLC
8500 Normandale Lake Blvd., Suite 320
Minneapolis, MN 55437
952.854.2700

Date: September 17, 2010

By: /Erin Nichols Matkaiti/

Erin Nichols Matkaiti
Reg. No. 57,125

PTO/SB/22 (07-09)
Approved for use through 07/31/2012. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARMENT OF COMMERCE
Under the paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

PETITION FOR EXTENSION OF TIME UNDER 37 CFR 1.136(a) **FY 2009** (*Fees pursuant to the Consolidated Appropriations Act, 2005 (H.R. 4818).*)	Docket Number (Optional) TPO.004.C2
Application Number 11/338,934	Filed January 25, 2006
For Grouping of Image Frames in Video Coding	
Art Unit 2611	Examiner Lee, S.

This is a request under the provisions of 37 CFR 1.136(a) to extend the period for filing a reply in the above identified application.

The requested extension and fee are as follows (check time period desired and enter the appropriate fee below):

		Fee	Small Entity Fee	
☑	One month (37 CFR 1.17(a)(1))	$130	$65	$ 130.00
☐	Two months (37 CFR 1.17(a)(2))	$490	$245	$
☐	Three months (37 CFR 1.17(a)(3))	$1110	$555	$
☐	Four months (37 CFR 1.17(a)(4))	$1730	$865	$
☐	Five months (37 CFR 1.17(a)(5))	$2350	$1175	$

☐ Applicant claims small entity status. See 37 CFR 1.27.

☐ A check in the amount of the fee is enclosed.

☐ Payment by credit card. Form PTO-2038 is attached.

☐ The Director has already been authorized to charge fees in this application to a Deposit Account.

☑ The Director is hereby authorized to charge any fees which may be required, or credit any overpayment, to Deposit Account Number 50-3581.

WARNING: Information on this form may become public. Credit card information should not be included on this form. Provide credit card information and authorization on PTO-2038.

I am the

☐ applicant/inventor.

☐ assignee of record of the entire interest. See 37 CFR 3.71.
Statement under 37 CFR 3.73(b) is enclosed (Form PTO/SB/96).

☑ attorney or agent of record. Registration Number 57,125

☐ attorney or agent under 37 CFR 1.34.
Registration number if acting under 37 CFR 1.34 _____

/Erin Nichols Matkaiti/ — Signature

September 17, 2010 — Date

Erin Nichols Matkaiti — Typed or printed name

952-854-2700 — Telephone Number

NOTE: Signatures of all the inventors or assignees of record of the entire interest or their representative(s) are required. Submit multiple forms if more than one signature is required, see below.

☑ Total of 1 forms are submitted.

This collection of information is required by 37 CFR 1.136(a). The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 6 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether disclosure of these records is required by the Freedom of Information Act.
2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.
3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.
4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).
5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.
6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).
7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.
8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspection or an issued patent.
9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

Electronic Patent Application Fee Transmittal

Application Number:	11338934
Filing Date:	25-Jan-2006
Title of Invention:	Grouping of image frames in video coding
First Named Inventor/Applicant Name:	Miska Hannuksela
Filer:	Erin M. Nichols
Attorney Docket Number:	TPO.004.C2

Filed as Large Entity

Utility under 35 USC 111(a) Filing Fees

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Basic Filing:				
Pages:				
Claims:				
Miscellaneous-Filing:				
Petition:				
Patent-Appeals-and-Interference:				
Post-Allowance-and-Post-Issuance:				
Extension-of-Time:				
Extension - 1 month with $0 paid	1251	1	130	130

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Miscellaneous:				
			Total in USD ($)	**130**

Electronic Acknowledgement Receipt

EFS ID:	8440165
Application Number:	11338934
International Application Number:	
Confirmation Number:	6610
Title of Invention:	Grouping of image frames in video coding
First Named Inventor/Applicant Name:	Miska Hannuksela
Customer Number:	76385
Filer:	Erin M. Nichols
Filer Authorized By:	
Attorney Docket Number:	TPO.004.C2
Receipt Date:	17-SEP-2010
Filing Date:	25-JAN-2006
Time Stamp:	11:49:56
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	yes
Payment Type	Deposit Account
Payment was successfully received in RAM	$130
RAM confirmation Number	10189
Deposit Account	503581
Authorized User	

The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows:

Charge any Additional Fees required under 37 C.F.R. Section 1.17 (Patent application and reexamination processing fees)

Charge any Additional Fees required under 37 C.F.R. Section 1.19 (Miscellaneous fees and charges)

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1		OAR-TPO004C2-OA25052010.pdf	95492 e2dea1ff0412ff322887c375b928c5a16e435a48	yes	10

Multipart Description/PDF files in .zip description		
Document Description	**Start**	**End**
Amendment/Req. Reconsideration-After Non-Final Reject	1	1
Claims	2	5
Applicant Arguments/Remarks Made in an Amendment	6	10

Warnings:

Information:

2	Extension of Time	ExtofTime-TPO004C2-OA25052010.pdf	312721 4d7d29170f3e25130fbb26d23c6e62c23ed6f1f3	no	2

Warnings:

Information:

3	Fee Worksheet (PTO-875)	fee-info.pdf	30074 f41cf99dda9e5d97013d2a985f9901328a3e8d1f	no	2

Warnings:

Information:

Total Files Size (in bytes):	438287

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111

If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371

If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office

If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

PTO/SB/06 (07-06)
Approved for use through 1/31/2007. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

PATENT APPLICATION FEE DETERMINATION RECORD

Substitute for Form PTO-875

Application or Docket Number	Filing Date	☐ To be Mailed
11/338,934	01/25/2006	

APPLICATION AS FILED – PART I

FOR	(Column 1) NUMBER FILED	(Column 2) NUMBER EXTRA	SMALL ENTITY ☐ RATE ($)	SMALL ENTITY ☐ FEE ($)	OR	OTHER THAN SMALL ENTITY RATE ($)	OTHER THAN SMALL ENTITY FEE ($)
☐ BASIC FEE (37 CFR 1.16(a), (b), or (c))	N/A	N/A	N/A			N/A	
☐ SEARCH FEE (37 CFR 1.16(k), (i), or (m))	N/A	N/A	N/A			N/A	
☐ EXAMINATION FEE (37 CFR 1.16(o), (p), or (q))	N/A	N/A	N/A			N/A	
TOTAL CLAIMS (37 CFR 1.16(i))	minus 20 =	*	X $ =		OR	X $ =	
INDEPENDENT CLAIMS (37 CFR 1.16(h))	minus 3 =	*	X $ =			X $ =	
☐ APPLICATION SIZE FEE (37 CFR 1.16(s))	If the specification and drawings exceed 100 sheets of paper, the application size fee due is $250 ($125 for small entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s).						
☐ MULTIPLE DEPENDENT CLAIM PRESENT (37 CFR 1.16(j))							
* If the difference in column 1 is less than zero, enter "0" in column 2.			TOTAL			TOTAL	

APPLICATION AS AMENDED – PART II

AMENDMENT 09/17/2010	(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA	SMALL ENTITY RATE ($)	SMALL ENTITY ADDITIONAL FEE ($)	OR	OTHER THAN SMALL ENTITY RATE ($)	OTHER THAN SMALL ENTITY ADDITIONAL FEE ($)
Total (37 CFR 1.16(i))	* 11	Minus	** 20	=	X $ =		OR	X $ =	
Independent (37 CFR 1.16(h))	* 6	Minus	***6	=	X $ =		OR	X $ =	
☐ Application Size Fee (37 CFR 1.16(s))									
☐ FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))							OR		
					TOTAL ADD'L FEE		OR	TOTAL ADD'L FEE	

AMENDMENT	(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA	RATE ($)	ADDITIONAL FEE ($)		RATE ($)	ADDITIONAL FEE ($)
Total (37 CFR 1.16(i))	*	Minus	**	=	X $ =		OR	X $ =	
Independent (37 CFR 1.16(h))	*	Minus	***	=	X $ =		OR	X $ =	
☐ Application Size Fee (37 CFR 1.16(s))									
☐ FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))							OR		
					TOTAL ADD'L FEE		OR	TOTAL ADD'L FEE	

* If the entry in column 1 is less than the entry in column 2, write "0" in column 3.

** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 20, enter "20".

*** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 3, enter "3".

The "Highest Number Previously Paid For" (Total or Independent) is the highest number found in the appropriate box in column 1.

Legal Instrument Examiner:
/PEGGY YARBOROUGH/

This collection of information is required by 37 CFR 1.16. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

FORM 1449* INFORMATION DISCLOSURE STATEMENT IN AN APPLICATION (Use several sheets if necessary)	Docket Number: TPO.004.C2	Application Number: 11/338,934
	Applicant: HANNUKSELA	
	Filing Date: 1/25/2006	Group Art Unit: 2611

U.S. PATENT DOCUMENTS

EXAMINER INITIAL	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE

FOREIGN PATENT DOCUMENTS

	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	TRANSLATION	
						YES	NO

OTHER DOCUMENTS (Including Author, Title, Date, Pertinent Pages, Etc.)

	9/15/2010	Office Action Response dated 9/15/2010 from U.S. Serial No. 10/306,942, 12 pages.
	9/10/2010	Office Action Response dated 9/10/2010 from U.S. Serial No. 11/338,996, 10 pages.

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609; draw line through citation if not in conformance and not considered. Include copy of this form for next communication to the Applicant.

*Substitute Disclosure Statement Form (PTO-1449) Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

IN THE UNITED STATES PATENT & TRADEMARK OFFICE

Applicant:	HANNUKSELA	Examiner:	Lee, S.
Serial No.:	11/338,934	Group Art Unit:	2611
Filed:	January 25, 2006	Docket No.:	TPO.004.C2
Title:	GROUPING OF IMAGE FRAMES IN VIDEO CODING		

SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT
UNDER 37 C.F.R. § 1.97(c)

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

Applicant respectfully submits the items of information on the enclosed Form 1449 for the attention of the Examiner in the above-identified application.

This statement should be considered because it is submitted after the mailing date of the first Office action on the merits but before the mailing date of a final Office action under 37 C.F.R. § 1.113 or a Notice of Allowance under 37 C.F.R. § 1.311. Please charge Deposit Account 50-3581 (TPO.004.C2) the amount of $180.00 under 37 C.F.R. § 1.17(p) for consideration of the items.

In accordance with 37 C.F.R. § 1.98(a)(2), only a copy of each foreign document or non-U.S. patent/application listed on the enclosed Form 1449 is provided.

Please note that any notations or markings on the attached documents do not reflect particular relevance, or lack thereof, to the present application, nor were they necessarily made by anyone affiliated with the prosecution of the present application.

No representation is made that a reference is "prior art" within the meaning of 35 U.S.C. §§ 102 and 103. In addition, Applicant does not represent that a reference has been thoroughly reviewed or that any relevance of any portion of a reference is intended, and reserves the right to establish otherwise under 37 C.F.R. § 1.131 or others.

Consideration of the items listed is respectfully requested. According to MPEP § 609, Applicant requests that the Examiner return a copy of the attached Form 1449, marked as being considered and initialed by the Examiner, to the undersigned with the next official communication.

Respectfully submitted,

HOLLINGSWORTH & FUNK, LLC
8500 Normandale Lake Blvd., Suite 320
Minneapolis, MN 55437
952.854.2700

By: /Erin Nichols Matkaiti/
Erin Nichols Matkaiti
Reg. No. 57,125

Electronic Patent Application Fee Transmittal

Application Number:	11338934
Filing Date:	25-Jan-2006
Title of Invention:	Grouping of image frames in video coding
First Named Inventor/Applicant Name:	Miska Hannuksela
Filer:	Erin M. Nichols
Attorney Docket Number:	TPO.004.C2

Filed as Large Entity

Utility under 35 USC 111(a) Filing Fees

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Basic Filing:				
Pages:				
Claims:				
Miscellaneous-Filing:				
Petition:				
Patent-Appeals-and-Interference:				
Post-Allowance-and-Post-Issuance:				
Extension-of-Time:				

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Miscellaneous:				
Submission- Information Disclosure Stmt	1806	1	180	180
		Total in USD ($)		**180**

Electronic Acknowledgement Receipt

EFS ID:	8531082
Application Number:	11338934
International Application Number:	
Confirmation Number:	6610
Title of Invention:	Grouping of image frames in video coding
First Named Inventor/Applicant Name:	Miska Hannuksela
Customer Number:	76385
Filer:	Erin M. Nichols
Filer Authorized By:	
Attorney Docket Number:	TPO.004.C2
Receipt Date:	30-SEP-2010
Filing Date:	25-JAN-2006
Time Stamp:	15:19:31
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	yes
Payment Type	Deposit Account
Payment was successfully received in RAM	$180
RAM confirmation Number	4605
Deposit Account	503581
Authorized User	

The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows:

Charge any Additional Fees required under 37 C.F.R. Section 1.17 (Patent application and reexamination processing fees)

Charge any Additional Fees required under 37 C.F.R. Section 1.19 (Miscellaneous fees and charges)

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Information Disclosure Statement (IDS) Filed (SB/08)	1449-11-TPO004C2.pdf	51214 1fa897d699f527cf94be8e8ae5e3eb24ec21e7b5	no	1
Warnings:					
Information:					
This is not an USPTO supplied IDS fillable form					
2	NPL Documents	OAR-TAPA004PA-OA23062010.pdf	97281 8e13d6c942b64fe9bc59ac82106b7e9c0b09ce2a	no	12
Warnings:					
Information:					
3	NPL Documents	OAR-TPO004C1-OA17062010.pdf	89215 26aba7cc752c3fc6ba6ddd2ef825b135aa56074e	no	10
Warnings:					
Information:					
4	Transmittal Letter	IDS-11-TPO004C2.pdf	55429 ca0531d1d8e8565dfc486c03a855f3ba7796a8f1	no	2
Warnings:					
Information:					
5	Fee Worksheet (PTO-875)	fee-info.pdf	30218 c426fea1268d247a4eb953a81801d64fed4fdb26	no	2
Warnings:					
Information:					
		Total Files Size (in bytes):	323357		

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111

If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371

If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office

If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

FORM 1449* SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT IN AN APPLICATION (Use several sheets if necessary)	Docket Number: TPO.004.C2	Application Number: 11/338,934
	Applicant: HANNUKSELA	
	Filing Date: 01/25/2006	Group Art Unit: 2611

U.S. PATENT DOCUMENTS

EXAMINER INITIAL	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE

FOREIGN PATENT DOCUMENTS

	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	TRANSLATION	
						YES	NO
	JP11275576	10/08/1999	JP			Abstract only	
	JP4348690	12/03/1992	JP			Abstract only	

OTHER DOCUMENTS (Including Author, Title, Date, Pertinent Pages, Etc.)

	08/17/2010	Copy of translated Office Action dated 08/17/2010 from parallel Japanese Application No. 2006-120297, 4 pages.
	08/24/2010	Copy of translated Office Action dated 08/24/2010 from parallel Japanese Application No. 2006-120296, 5 pages.

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609; draw line through citation if not in conformance and not considered. Include copy of this form for next communication to the Applicant.

*Substitute Disclosure Statement Form (PTO-1449) Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

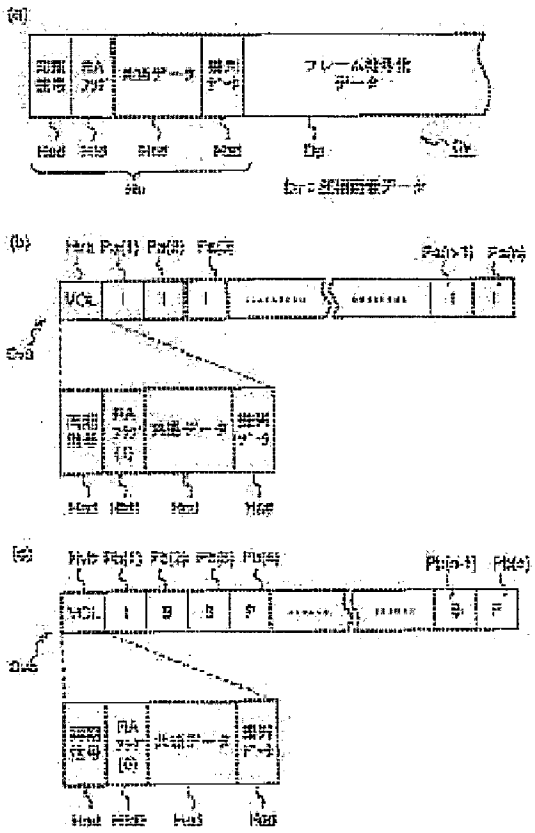
IMAGE TRANSMISSION METHOD, IMAGE PROCESSING METHOD, IMAGE PROCESSOR AND DATA STORAGE MEDIUM

Publication number: JP11275576 (A)
Publication date: 1999-10-08
Inventor(s): BUN CHUN SEN +
Applicant(s): MATSUSHITA ELECTRIC IND CO LTD +
Classification:
- international: ***H04N7/24; H04N7/26; H04N7/32;*** **H04N7/24; H04N7/26; H04N7/32;** (IPC1-7): H04N7/24
- European:
Application number: JP19990014735 19990122
Priority number(s): JP19990014735 19990122; JP19980011068 19980123

Also published as:
JP3380763 (B2)

Abstract of **JP 11275576 (A)**

PROBLEM TO BE SOLVED: To detect aptitude of random independent reproduction processing in a short time with respect to compressed image data through the analysis of an identification flag in the case of analysis processing of a header imparted to the compressed image data · corresponding to a moving image.
SOLUTION: In the image transmission method, compressed image data Dv obtained by applying compression coding processing to digital image data corresponding to a moving image are transmitted, and in this case, an identification flag Hfd denoting whether or not the compressed image data Dv are proper to random independent reproduction onto an optional screen is sent next to a synchronizing signal Hsd placed at a head of a header Hv.



Data supplied from the ***espacenet*** database — Worldwide

(19)日本国特許庁（ＪＰ）　(12) **公開特許公報**（Ａ）　(11)特許出願公開番号

特開平11－275576

(43)公開日　平成11年(1999)10月８日

(51)Int.Cl.[6]	識別記号	ＦＩ	
Ｈ０４Ｎ　7/24		Ｈ０４Ｎ　7/13	Ｚ

審査請求　未請求　請求項の数22　ＯＬ　（全 31 頁）

(21)出願番号　特願平11－14735

(22)出願日　平成11年(1999)１月22日

(31)優先権主張番号　特願平10－11068

(32)優先日　平10(1998)１月23日

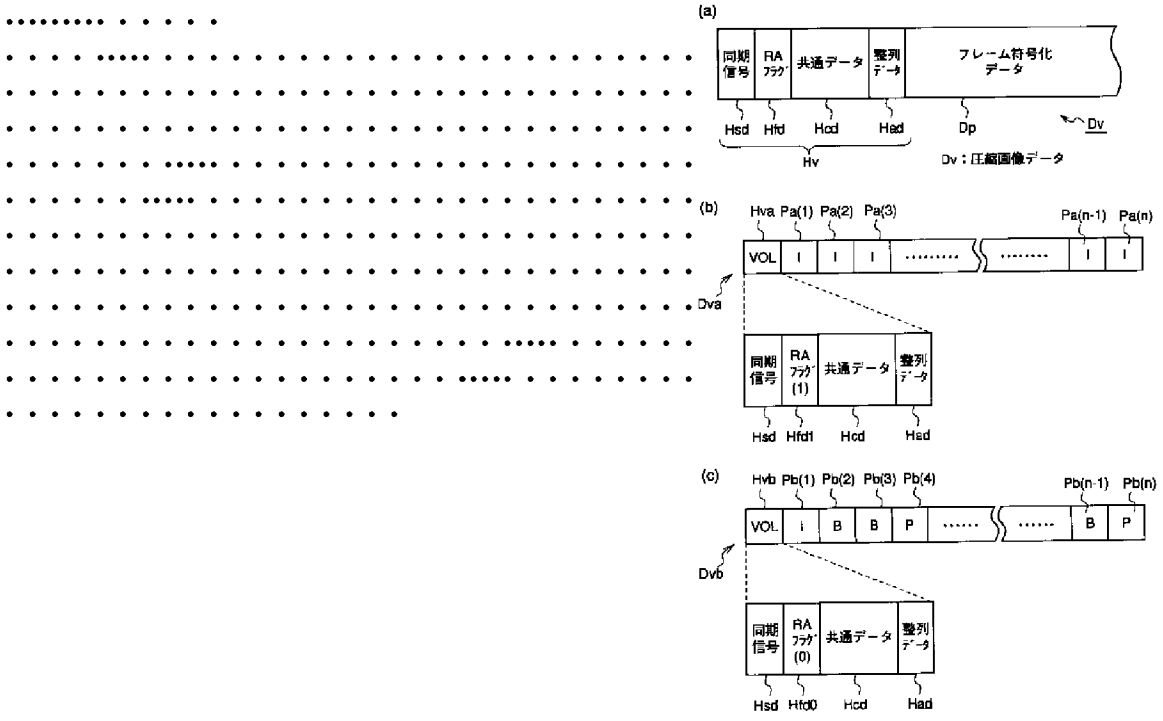
(33)優先権主張国　日本（ＪＰ）

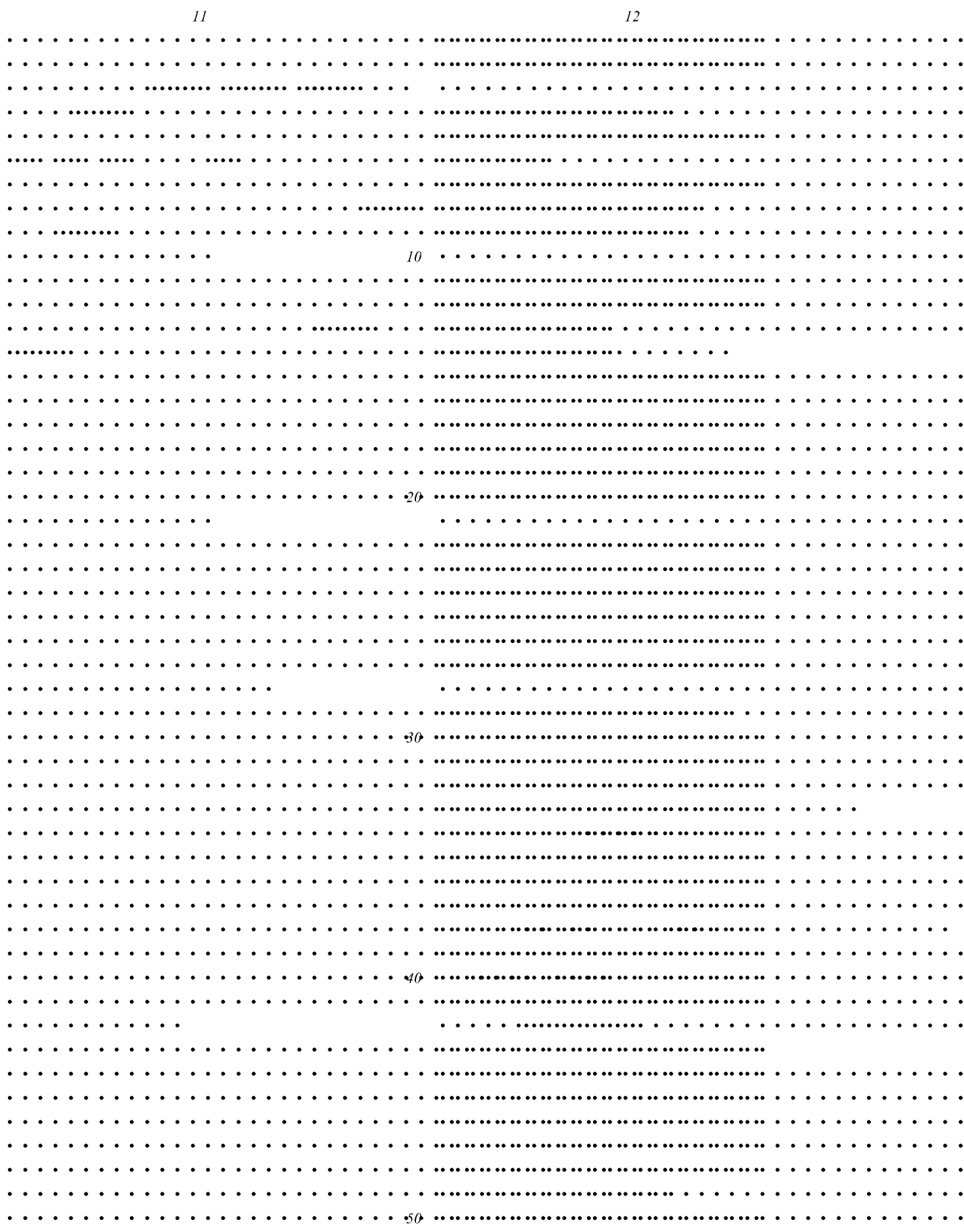
(71)出願人　000005821
松下電器産業株式会社
大阪府門真市大字門真1006番地

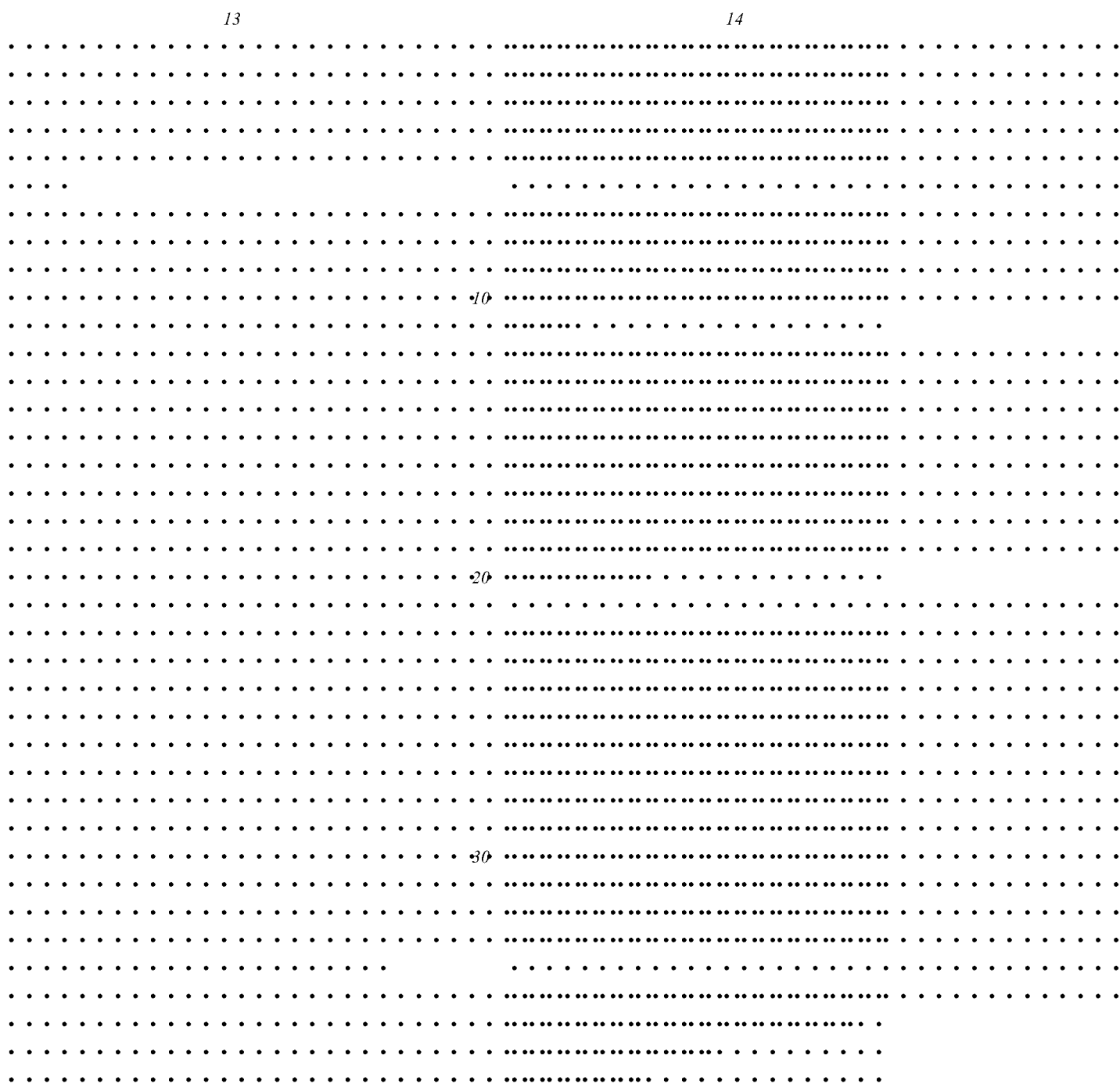
(72)発明者　ブン　チュン　セン
大阪府門真市大字門真1006番地　松下電器産業株式会社内

(74)代理人　弁理士　早瀬　憲一

(54)【発明の名称】　画像伝送方法，画像処理方法，画像処理装置，及びデータ記憶媒体



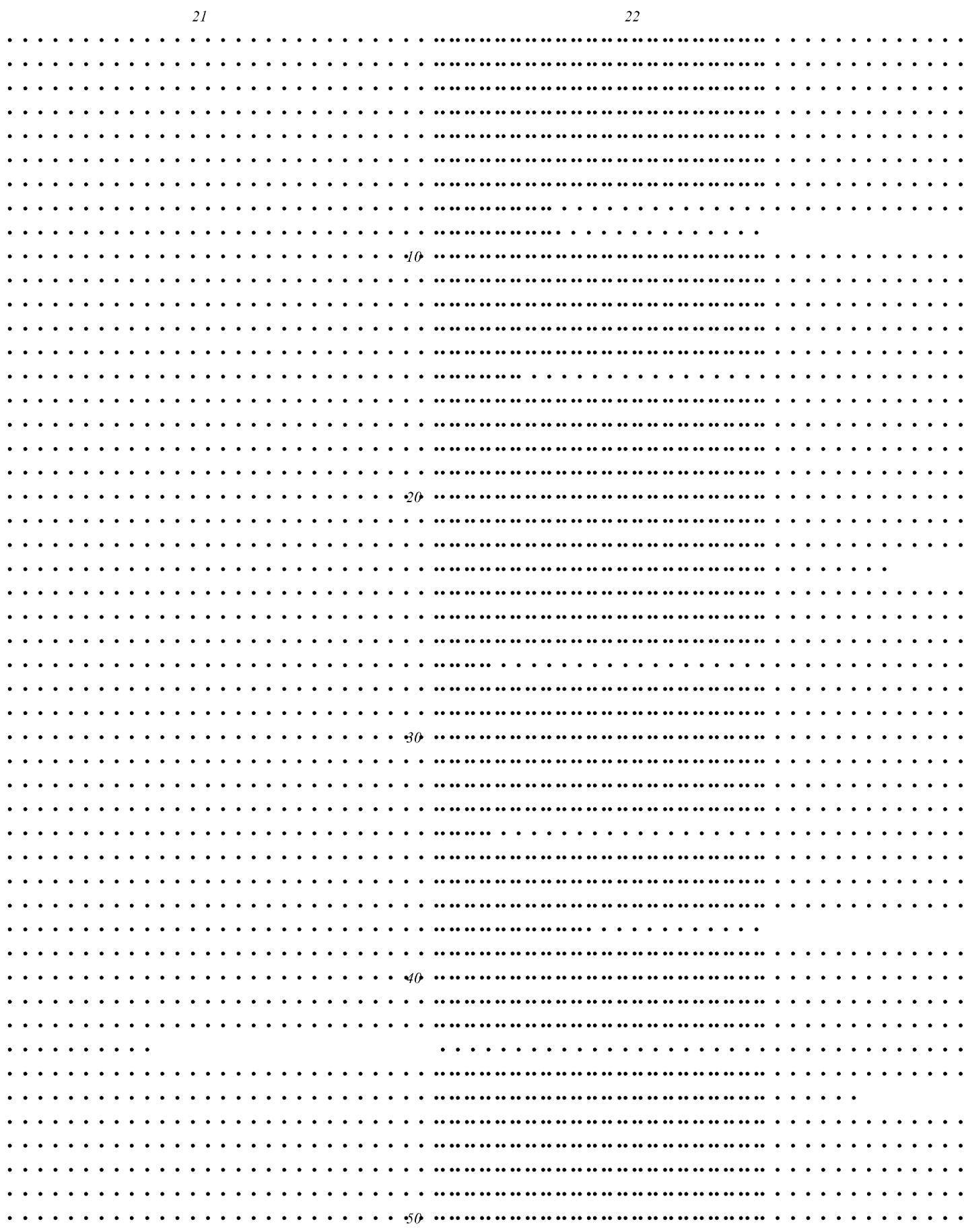

11
12
10
20
30
40
50

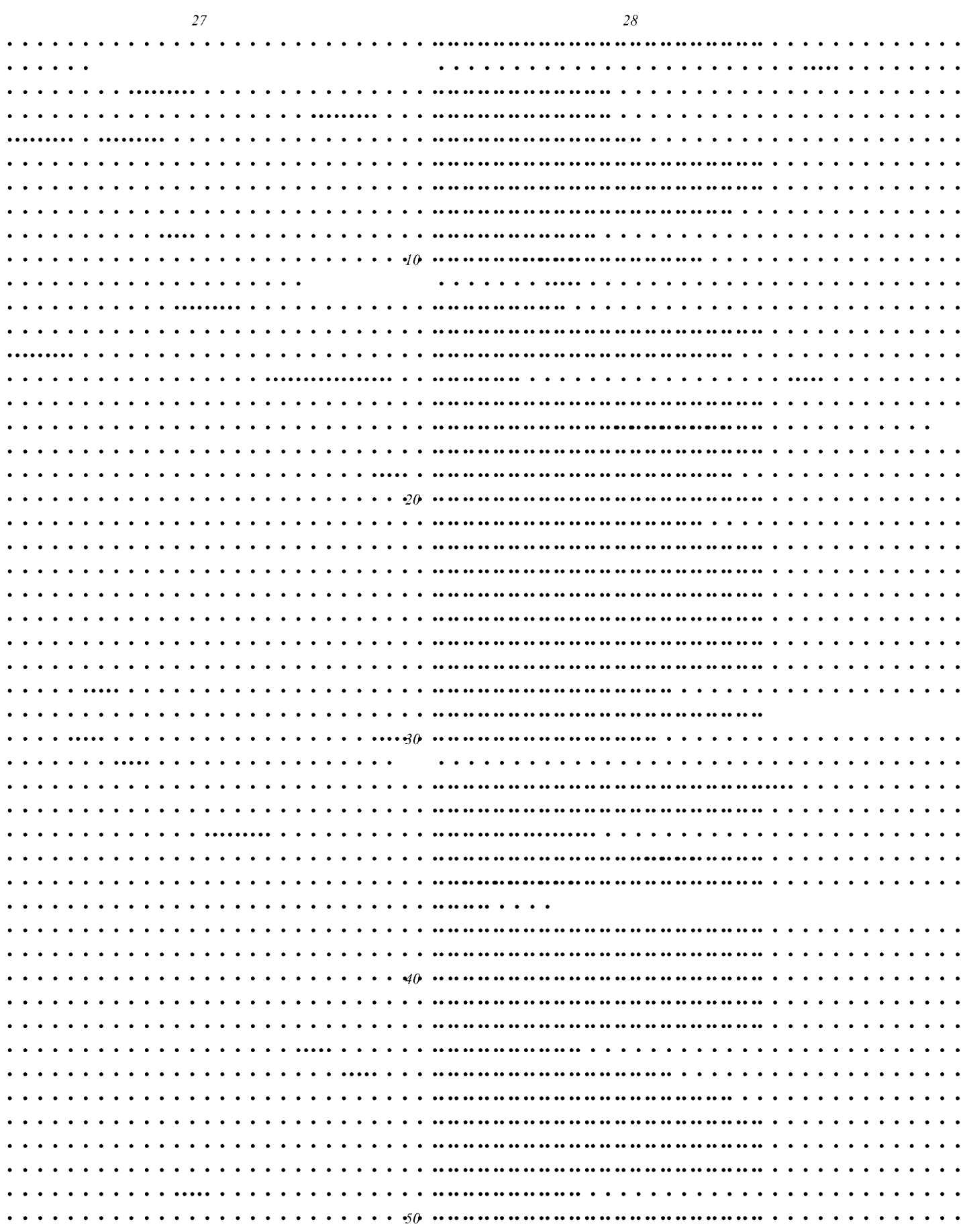

13
14
10
20
30

		No.of bits	Mnemonic
901	Video Object Layer(){		
902	video_object_layer_start_code /*4least significant bits specify video_object_layer_id value*/	32	bslbf
903	is_object_layer_identifier ~903a	1	uimsbf
	if(is_object_layer_identifier){		
	visual_object_layer_verid ~903b	4	uimsbf
	visual_object_layer_priority ~903c	3	uimsbf
	}		
904	vol_control_parameters	1	bslbf
	if(vol_control_parameters)		
	aspect_ratio_info	4	uimsbf
	vop_rate_code	4	uimsbf
	bit_rate	30	uimsbf
	vbv_buffer_size	18	uimsbf
	chroma_format	2	uimsbf
	low_delay	1	uimsbf
	}		
905	video_object_layer_shape	2	uimsbf
	vop_time_increment_resolution	15	uimsbf
	fixed_vop_rate	1	bslbf
906	if(video_object_layer_shape!="binary only"){		
907	if(video_object_layer_shape=="rectangular"){		
	marker_bit	1	bslbf
	video_object_layer_width	13	uimsbf
	marker_bit	1	bslbf
	video_object_layer_height	13	uimsbf
	}		
	obmc_disable	1	bslbf

908	sprite_enable	1	bslbf
	if(sprite_enable){		
	sprite_width	13	uimsbf
	marker_bit	1	bslbf
	sprite_height	13	uimsbf
	marker_bit	1	bslbf
	sprite_left_coodinate	13	simsbf
	marker_bit	1	bslbf
	sprite_top_coodinate	13	simsbf
	marker_bit	1	bslbf
	no_of_sprite_warping_points	6	uimsbf
	sprite_warping_accuracy	2	uimsbf
	sprite_brightness_change	1	bslbf
	if(video_object_layer_shape=="rectangular"){		
	init_sprite_width	13	uimsbf
	marker_bit	1	bslbf
	init_sprite_height	13	uimsbf
	marker_bit	1	bslbf
	init_sprite_left_coodinate	13	simsbf
	marker_bit	1	bslbf
	init_sprite_top_coodinate	13	simsbf
	}		
	}		
909	not_8_bit	1	bslbf
	if(not_8_bit){		
	quant_precision	4	uimsbf
	bits_per_pixel	4	uimsbf
	}		
910	quant_type	1	bslbf
	if(quant_type){		
	load_intra_quant_mat	1	bslbf
	if(load_intra_quant_mat)		
	intra_quant_mat	8*[2-64]	uimsbf
	load_nonintra_quant_mat	1	bslbf
	if(load_nonintra_quant_mat)		
	nonintra_quant_mat	8*[2-64]	uimsbf
	}		
911	complexity_estimation_disable	1	bslbf

912	error_resilient_disable	1	bslbf
	if(!error_resilient_disable){		
	data_partitioned	1	bslbf
	reversible_vlc	1	bslbf
	}		
913	scalability	1	bslbf
	if(scalability){		
	ref_layer_id	4	uimsbf
	ref_layer_sampling_direc	1	bslbf
	hor_sampling_factor_n	5	uimsbf
	hor_sampling_factor_m	5	uimsbf
	vert_sampling_factor_n	5	uimsbf
	vert_sampling_factor_m	5	uimsbf
	enhancement_type	1	bslbf
	}		
	}		
914	random_accessible_vol	1	bslbf
915	next_start_code()		
916	if(sprite_enable)		
	decode_init_sprite()		
917	do{		
	if(next_bits()=group_of_vop_start_code)		
	Group_of_Video object Plane()		
	Video Object Plane()		
	}while((next_bits()=group_of_vop_start_code)|| (next_bits()=vop_start_code)		
	}		


21
22
10
20
30
40
50

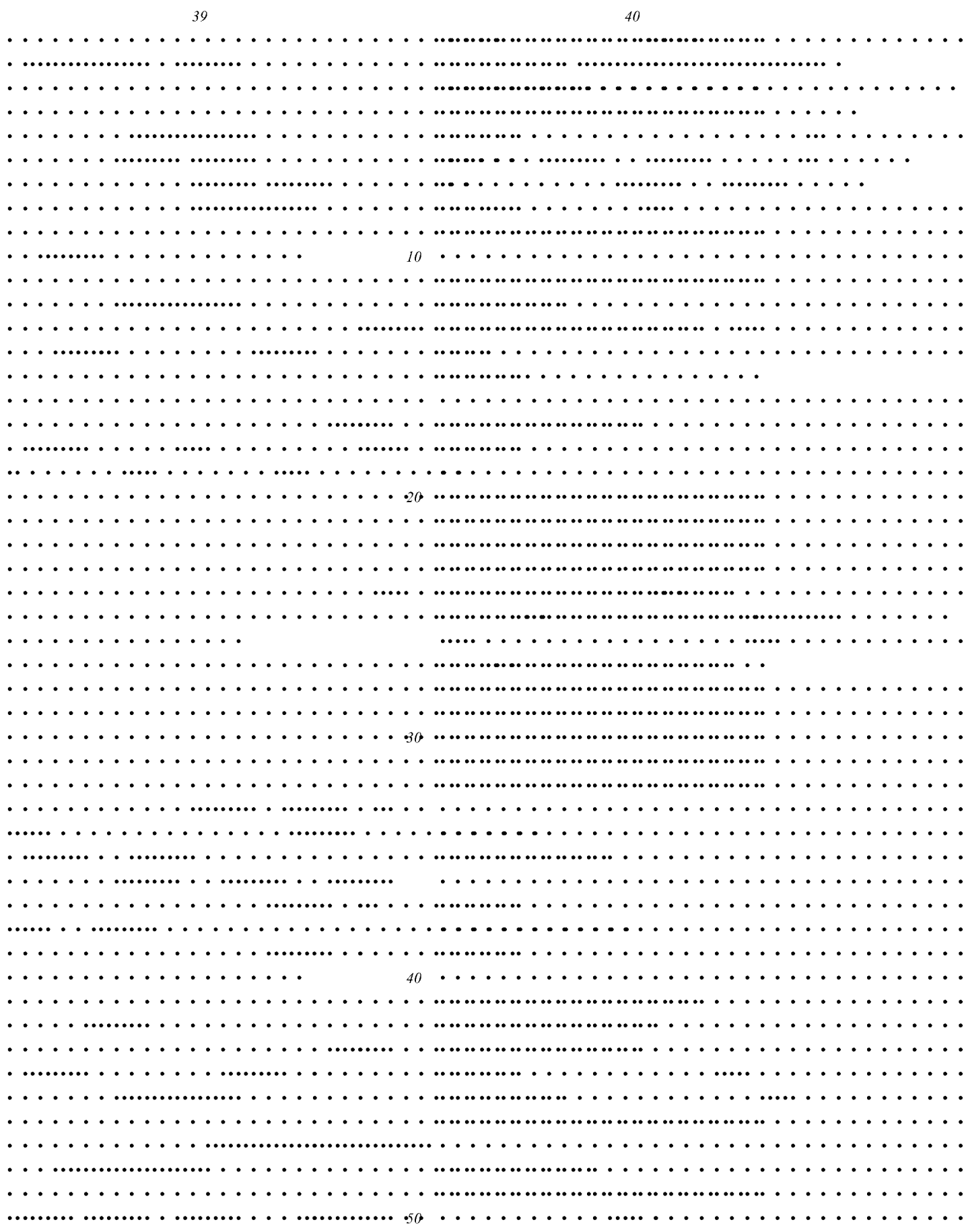


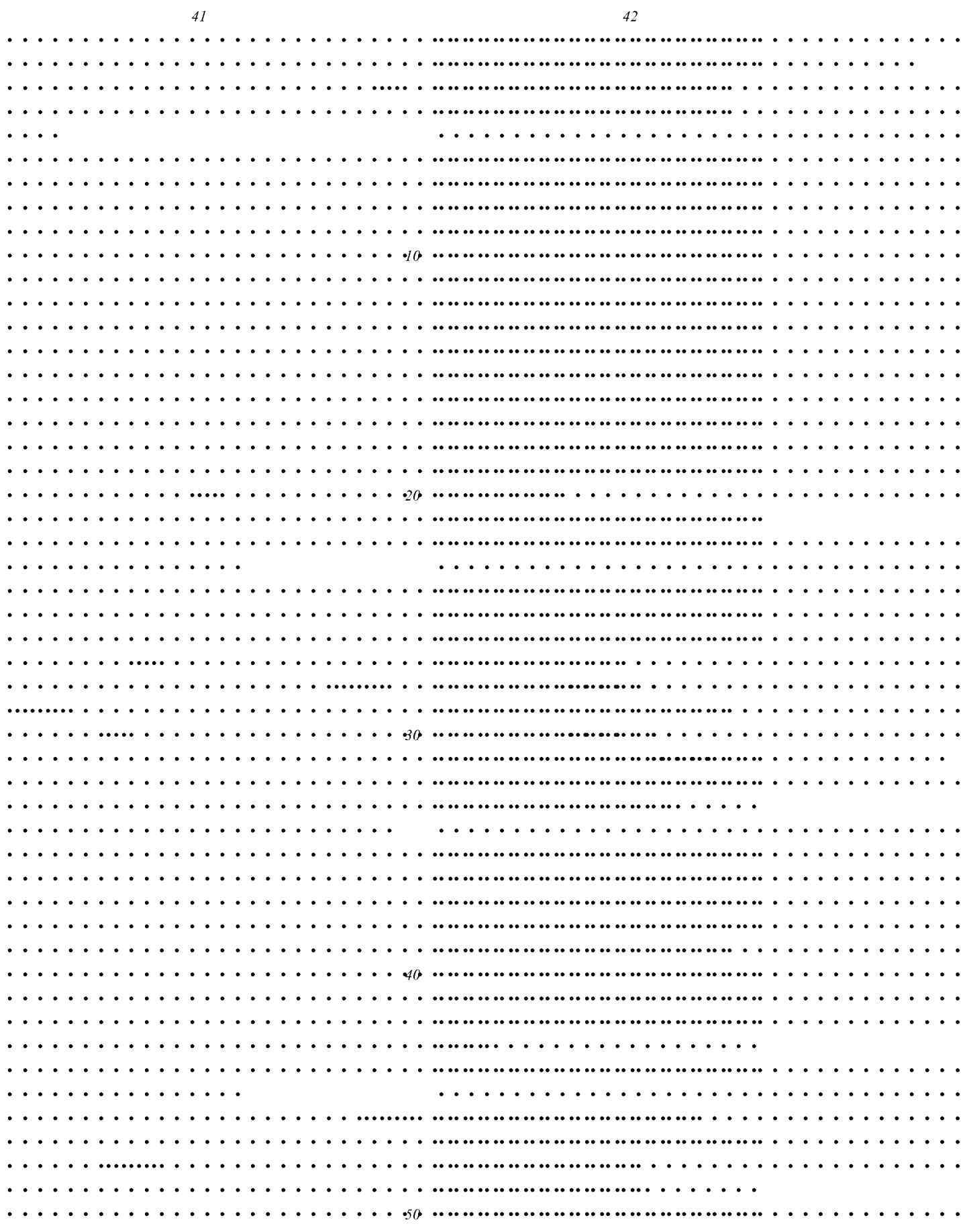
	Video Object Layer(){	No.of bits	Mnemonic
802	video_object_layer_start_code /*4least significant bits specify video_object_layer_id value*/	32	bslbf
814	random_accessible_vol	1	bslbf
803	is_object_layer_identifier	1	uimsbf
	if(is_object_layer_identifier){		
	visual_object_layer_verid	4	uimsbf
	visual_object_layer_priority	3	uimsbf
	}		
804	vol_control_parameters	1	bslbf
	if(vol_control_parameters)		
	aspect_ratio_info	4	uimsbf
	vop_rate_code	4	uimsbf
	bit_rate	30	uimsbf
	vbv_buffer_size	18	uimsbf
	chroma_format	2	uimsbf
	low_delay	1	uimsbf
	}		
805	video_object_layer_shape	2	uimsbf
	vop_time_increment_resolution	15	uimsbf
	fixed_vop_rate	1	bslbf
806	if(video_object_layer_shape!="binary only"){		
807	if(video_object_layer_shape=="rectangular"){		
	marker_bit	1	bslbf
	video_object_layer_width	13	uimsbf
	marker_bit	1	bslbf
	video_object_layer_height	13	uimsbf
	}		
	obmc_disable	1	bslbf

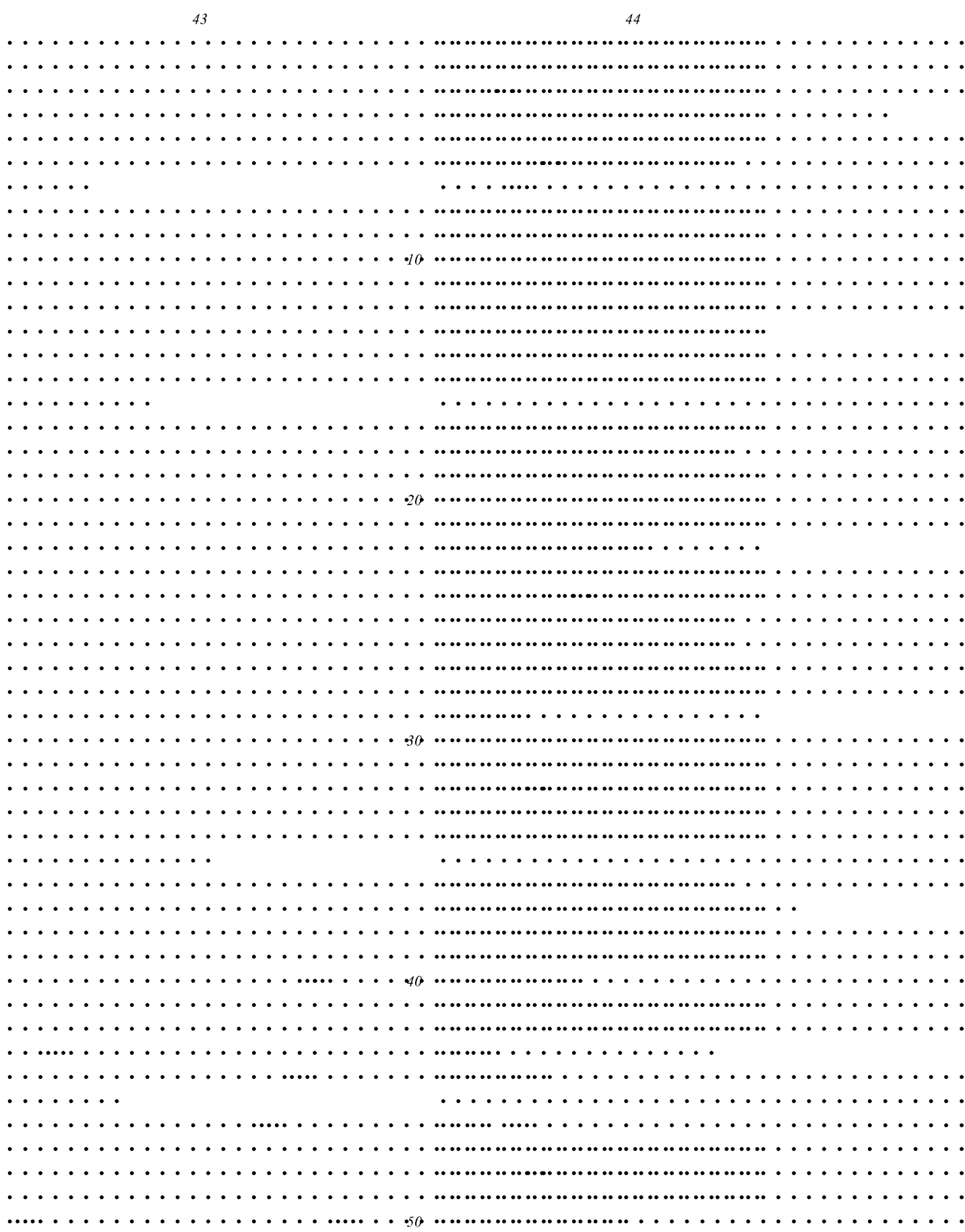
31 32

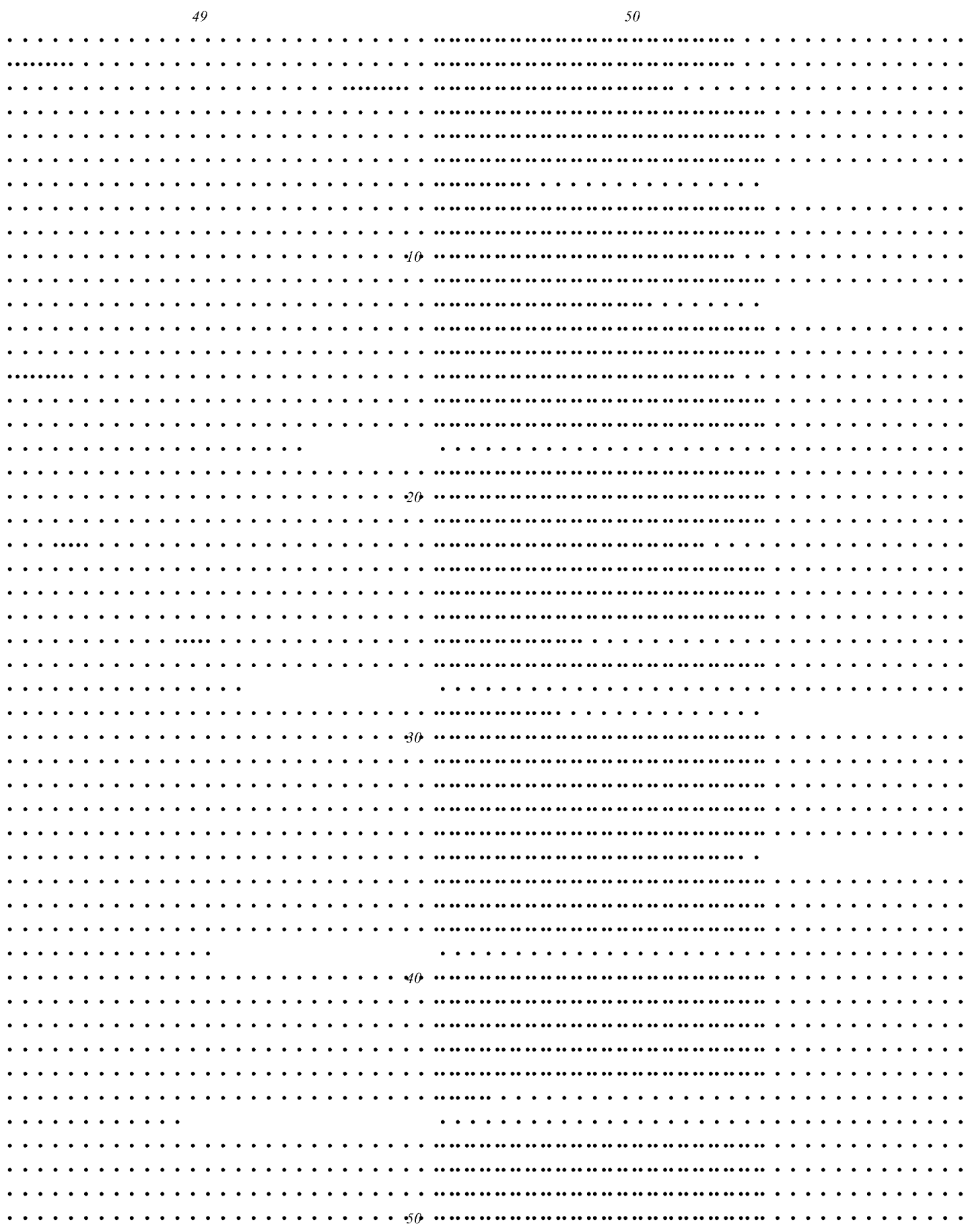
808	sprite_enable	1	bslbf
	if(sprite_enable){		
	sprite_width	13	uimsbf
	marker_bit	1	bslbf
	sprite_height	13	uimsbf
	marker_bit	1	bslbf
	sprite_left_coodinate	13	simsbf
	marker_bit	1	bslbf
	sprite_top_coodinate	13	simsbf
	marker_bit	1	bslbf
	no_of_sprite_warping_points	6	uimsbf
	sprite_warping_accuracy	2	uimsbf
	sprite_brightness_change	1	bslbf
	if(video_object_layer_shape=="rectangular"){		
	init_sprite_width	13	uimsbf
	marker_bit	1	bslbf
	init_sprite_height	13	uimsbf
	marker_bit	1	bslbf
	init_sprite_left_coodinate	13	simsbf
	marker_bit	1	bslbf
	init_sprite_top_coodinate	13	simsbf
	}		
	}		
809	not_8_bit	1	bslbf
	if(not_8_bit){		
	quant_precision	4	uimsbf
	bits_per_pixel	4	uimsbf
	}		
810	quant_type	1	bslbf
	if(quant_type){		
	load_intra_quant_mat	1	bslbf
	if(load_intra_quant_mat)		
	intra_quant_mat	8*[2-64]	uimsbf
	load_nonintra_quant_mat	1	bslbf
	if(load_nonintra_quant_mat)		
	nonintra_quant_mat	8*[2-64]	uimsbf
	}		
811	complexity_estimation_disable	1	bslbf

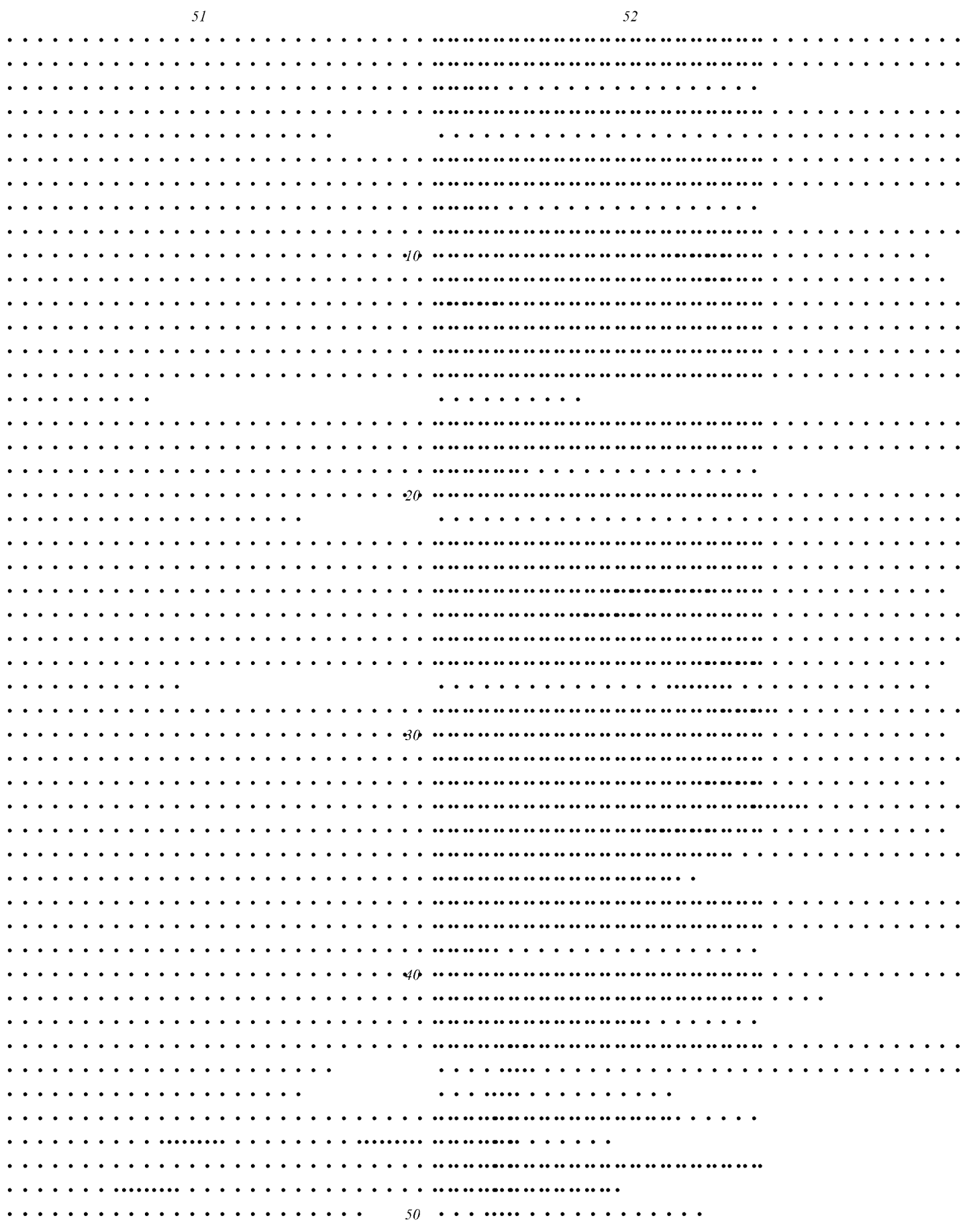
812	error_resilient_disable	1	bslbf
	if(!error_resilient_disable){		
	data_partitioned	1	bslbf
	reversible_vlc	1	bslbf
	}		
813	scalability	1	bslbf
	if(scalability){		
	ref_layer_id	4	uimsbf
	ref_layer_sampling_direc	1	bslbf
	hor_sampling_factor_n	5	uimsbf
	hor_sampling_factor_m	5	uimsbf
	vert_sampling_factor_n	5	uimsbf
	vert_sampling_factor_m	5	uimsbf
	enhancement_type	1	bslbf
	}		
	}		
815	next_start_code()		
816	if(sprite_enable)		
	decode_init_sprite()		
817	do{		
	if(next_bits()=group_of_vop_start_code)		
	Group_of_Video object Plane()		
	Video Object Plane()		
	}while((next_bits()=group_of_vop_start_code)|| (next_bits()=vop_start_code)		
	}		

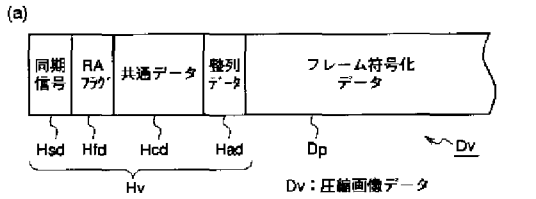

39
40
10
20
30
40
50

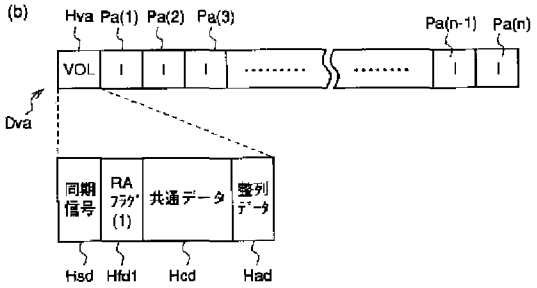

41
42
10
20
30
40
50

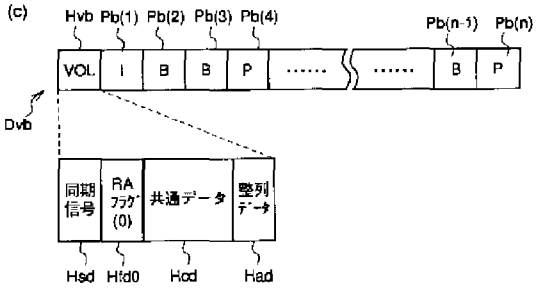


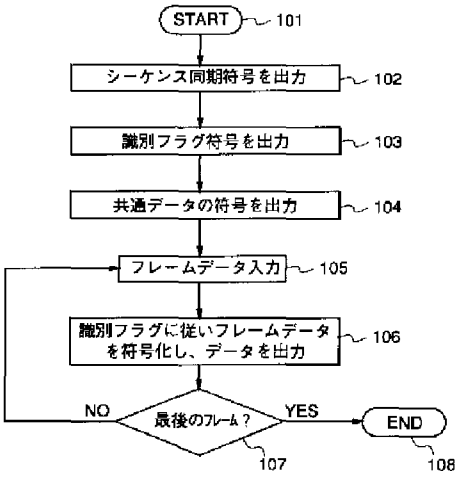

49
50
10
20
30
40
50

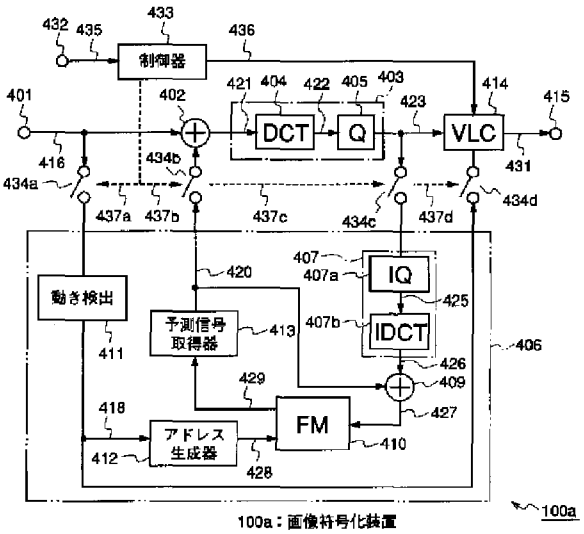

51
52
10
20
30
40
50

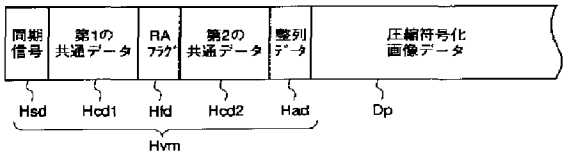

(a)
同期信号
RA フラグ
共通データ
整列データ
フレーム符号化データ
Hsd
Hfd
Hcd
Had
Dp
Dv
Hv
Dv：圧縮画像データ

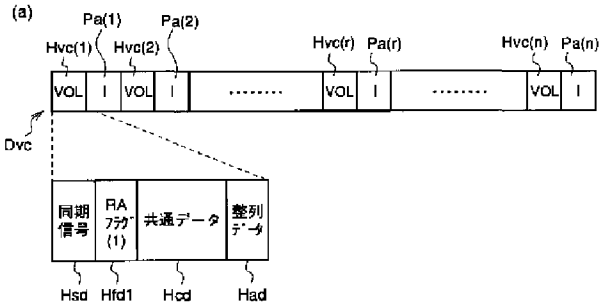

(b)
Hva Pa(1) Pa(2) Pa(3)
Pa(n-1) Pa(n)
VOL
I
Dva
同期信号
RA フラグ (1)
共通データ
整列データ
Hsd
Hfd1
Hcd
Had

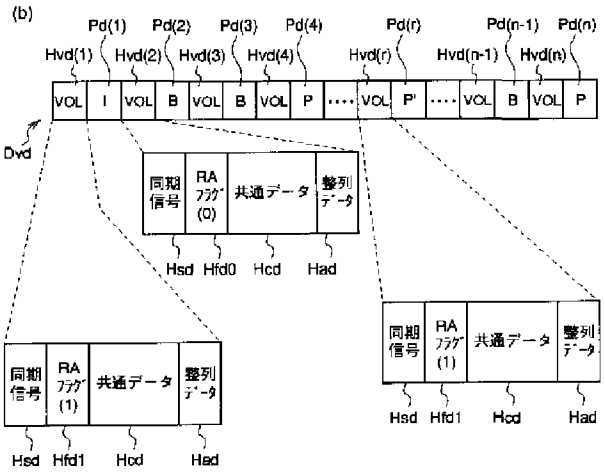

(c)
Hvb Pb(1) Pb(2) Pb(3) Pb(4)
Pb(n-1) Pb(n)
VOL
I
B
P
Dvb
同期信号
RA フラグ (0)
共通データ
整列データ
Hsd
Hfd0
Hcd
Had

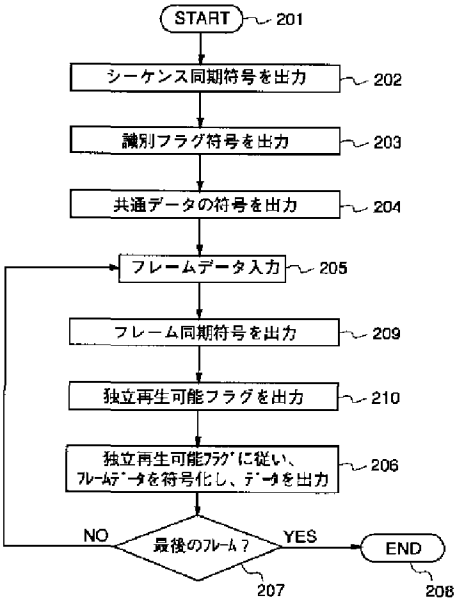

START
101
シーケンス同期符号を出力
102
識別フラグ符号を出力
103
共通データの符号を出力
104
フレームデータ入力
105
識別フラグに従いフレームデータを符号化し、データを出力
106
最後のフレーム？
NO
YES
107
END
108

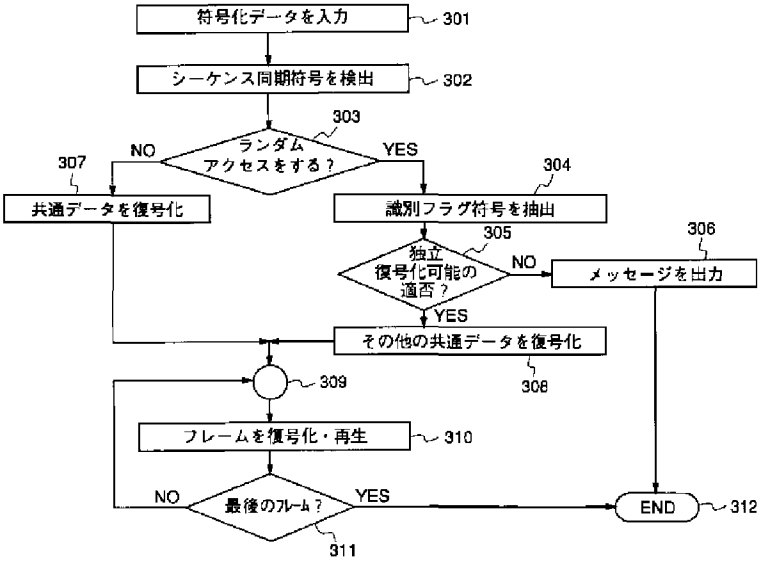

制御器
DCT
Q
VLC
IQ
IDCT
FM
動き検出
予測信号取得器
アドレス生成器
100a
100a：画像符号化装置

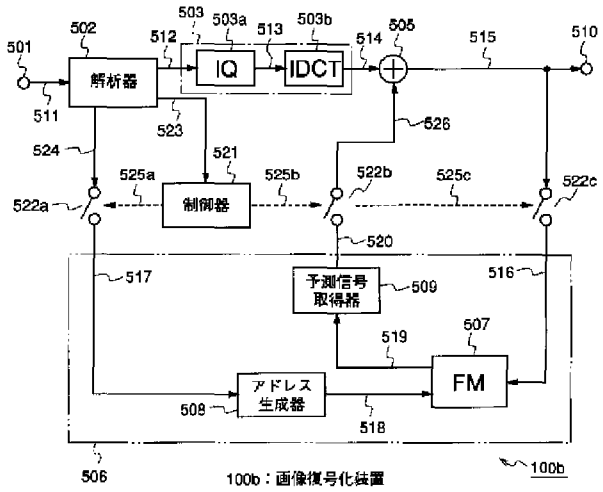

同期信号
第1の共通データ
RAフラグ
第2の共通データ
整列データ
圧縮符号化画像データ
Hsd
Hcd1
Hfd
Hcd2
Had
Dp
Hvm

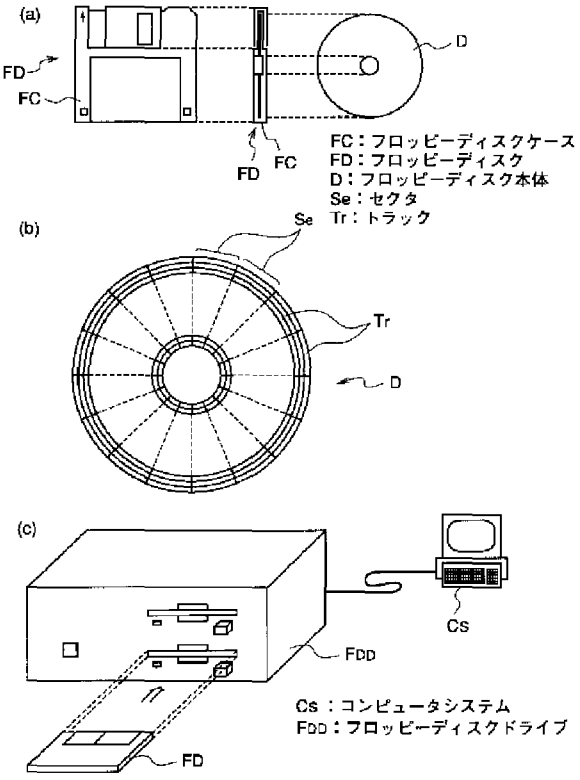

(a)
Pa(1)
Pa(2)
Hvc(1)
Hvc(2)
Hvc(r)
Pa(r)
Hvc(n)
Pa(n)
VOL
I
Dvc
同期信号
RAフラグ (1)
共通データ
整列データ
Hsd
Hfd1
Hcd
Had

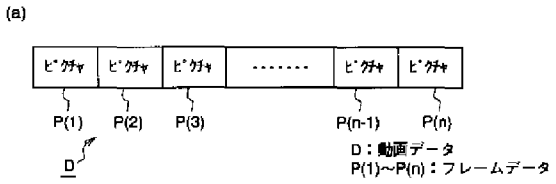

(b)
Pd(1)
Pd(2)
Pd(3)
Pd(4)
Pd(r)
Pd(n-1)
Pd(n)
Hvd(1)
Hvd(2)
Hvd(3)
Hvd(4)
Hvd(r)
Hvd(n-1)
Hvd(n)
VOL
I
B
P
P'
Dvd
同期信号
RAフラグ (0)
共通データ
整列データ
Hsd
Hfd0
Hcd
Had
RAフラグ (1)
Hfd1

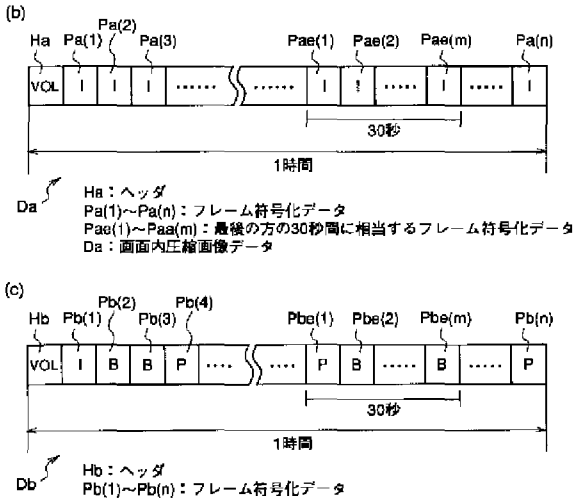

START
201
シーケンス同期符号を出力
202
識別フラグ符号を出力
203
共通データの符号を出力
204
フレームデータ入力
205
フレーム同期符号を出力
209
独立再生可能フラグを出力
210
独立再生可能フラグに従い、フレームデータを符号化し、データを出力
206
NO
最後のフレーム？
YES
207
END
208

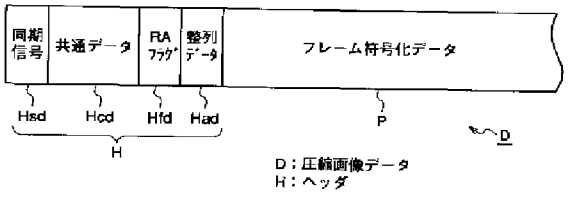

符号化データを入力
301
シーケンス同期符号を検出
302
303
ランダムアクセスをする？
NO
YES
307
304
共通データを復号化
識別フラグ符号を抽出
305
306
独立復号化可能の適否？
NO
メッセージを出力
YES
その他の共通データを復号化
308
309
フレームを復号化・再生
310
NO
最後のフレーム？
YES
311
END
312

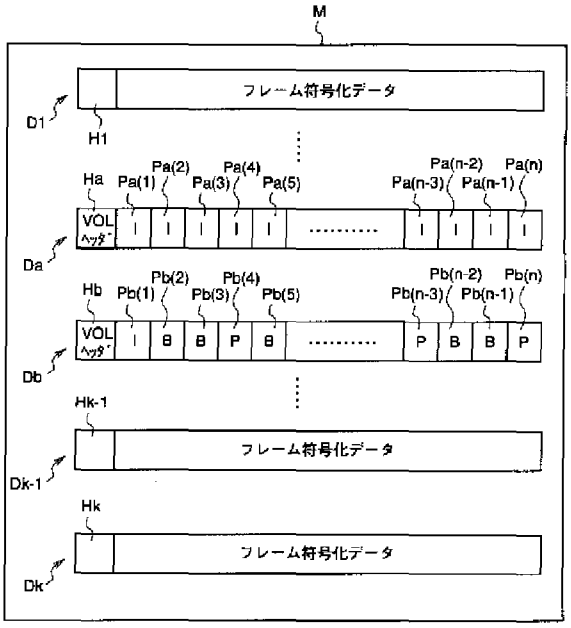

501
502
503
503a
503b
505
510
512
513
514
515
解析器
IQ
IDCT
511
523
526
524
521
525a
525b
522b
525c
522c
522a
制御器
520
517
516
予測信号
取得器
509
519
507
FM
アドレス
生成器
508
518
506
100b
100b：画像復号化装置

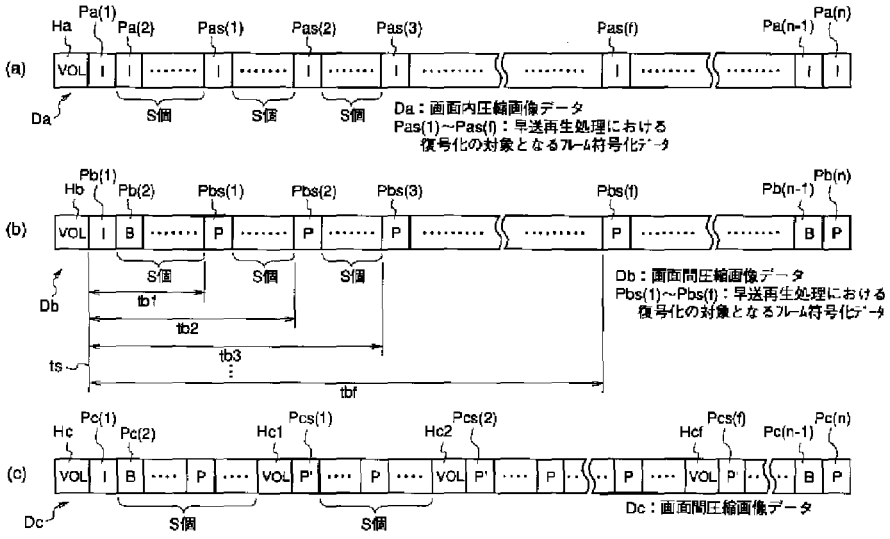

(a)
FD
FC
D
FC：フロッピーディスクケース
FD：フロッピーディスク
D：フロッピーディスク本体
Se：セクタ
Tr：トラック
(b)
Se
Tr
D
(c)
Cs
FDD
FD
Cs：コンピュータシステム
FDD：フロッピーディスクドライブ


(a)
ピクチャ
ピクチャ
ピクチャ
ピクチャ
ピクチャ
P(1)
P(2)
P(3)
P(n-1)
P(n)
D
D：動画データ
P(1)～P(n)：フレームデータ


(b)
Ha
Pa(1)
Pa(2)
Pa(3)
Pae(1)
Pae(2)
Pae(m)
Pa(n)
VOL
30秒
1時間
Da
Ha：ヘッダ
Pa(1)～Pa(n)：フレーム符号化データ
Pae(1)～Paa(m)：最後の方の30秒間に相当するフレーム符号化データ
Da：画面内圧縮画像データ
(c)
Hb
Pb(1)
Pb(2)
Pb(3)
Pb(4)
Pbe(1)
Pbe(2)
Pbe(m)
Pb(n)
VOL
I
B
P
30秒
1時間
Db
Hb：ヘッダ
Pb(1)～Pb(n)：フレーム符号化データ
Pbe(1)～Pbe(m)：最後の方の30秒間に相当するフレーム符号化データ
Db：画面間圧縮画像データ


同期信号
共通データ
RA フラグ
整列データ
フレーム符号化データ
Hsd
Hcd
Hfd
Had
H
P
D
D：圧縮画像データ
H：ヘッダ


M
D1
H1
フレーム符号化データ
Ha
Pa(1)
Pa(2)
Pa(3)
Pa(4)
Pa(5)
Pa(n-3)
Pa(n-2)
Pa(n-1)
Pa(n)
VOL
Da
Hb
Pb(1)
Pb(2)
Pb(3)
Pb(4)
Pb(5)
Pb(n-3)
Pb(n-2)
Pb(n-1)
Pb(n)
Db
Hk-1
Dk-1
Hk
Dk
M：データ記憶媒体


(a)
Ha
Pa(1)
Pa(2)
Pas(1)
Pas(2)
Pas(3)
Pas(f)
Pa(n-1)
Pa(n)
VOL
I
Da
S個
Da：画面内圧縮画像データ
Pas(1)～Pas(f)：早送再生処理における
復号化の対象となるフレーム符号化データ
(b)
Hb
Pb(1)
Pb(2)
Pbs(1)
Pbs(2)
Pbs(3)
Pbs(f)
Pb(n-1)
Pb(n)
B
P
Db
tb1
tb2
tb3
tbf
ts
Db：画面間圧縮画像データ
Pbs(1)～Pbs(f)：早送再生処理における
復号化の対象となるフレーム符号化データ
(c)
Hc
Pc(1)
Pc(2)
Hc1
Pcs(1)
Hc2
Pcs(2)
Hcf
Pcs(f)
Pc(n-1)
Pc(n)
VOL
I
B
P
P'
Dc
S個
Dc：画面間圧縮画像データ

PICTURE STORAGE TRANSMITTER

Publication number: JP4348690 (A)
Publication date: 1992-12-03
Inventor(s): TANAKA TOMOAKI; OBA YUJI; INOMATA HIDEKI +
Applicant(s): MITSUBISHI ELECTRIC CORP; NIPPON TELEGRAPH & TELEPHONE +
Classification:
- international: ***G06T3/40; G06T9/00; H04N1/41; H04N7/173; H04N7/24; H04N7/26;*** *H04L12/28;* **G06T3/40; G06T9/00; H04N1/41; H04N7/173; H04N7/24; H04N7/26;** H04L12/28; (IPC1-7): G06F15/66; H04L12/48; H04N7/13; H04N7/173
- European:
Application number: JP19910121148 19910527
Priority number(s): JP19910121148 19910527

Also published as:
JP2609772 (B2)

Abstract of **JP 4348690 (A)**

PURPOSE:To prevent mis-recognition as cell abort when a sequence number of cell between frames is discontinuous due to the frame edit of the cell in a picture data storage section. CONSTITUTION:After each cell with a sequence number added thereto in the ascending order is stored in a picture data storage section 2, when a storage control section 3 edits each cell in the unit of frames and the sequence numbers between frames outputted one after another are discontinuous, an identification cell multiplex section 4 is activated to multiplex the identification cell between the frames. As a result, an abort discrimination section 7 recognizes the identification cell to prevent the abort of a cell in the sequence number missing part from being mis-recognized regardless of the discontinuous sequence numbers.

Data supplied from the ***espacenet*** database — Worldwide

(19)日本国特許庁（ＪＰ）　(12)**公開特許公報**（Ａ）　(11)特許出願公開番号

特開平4－348690

(43)公開日　平成４年(1992)12月３日

(51)Int.Cl.[5]	識別記号	庁内整理番号	Ｆ Ｉ	技術表示箇所
Ｈ０４Ｎ　7/13	Ｚ	8838－5C		
Ｇ０６Ｆ　15/66	３５５　Ａ	8420－5L		
Ｈ０４Ｎ　7/173		8324－5C		
// Ｈ０４Ｌ　12/48				
		8529－5K	Ｈ０４Ｌ　11/ 20　Ｚ	

審査請求　未請求　請求項の数１(全　５　頁)

(21)出願番号　特願平3－121148

(22)出願日　平成３年(1991)５月27日

(71)出願人　000006013
三菱電機株式会社
東京都千代田区丸の内二丁目２番３号
(71)出願人　000004226
日本電信電話株式会社
東京都千代田区内幸町一丁目１番６号
(72)発明者　田中　知明
東京都千代田区内幸町一丁目１番６号　日本電信電話株式会社内
(72)発明者　大庭　有二
東京都千代田区内幸町一丁目１番６号　日本電信電話株式会社内
(74)代理人　弁理士　高田　守　(外２名)

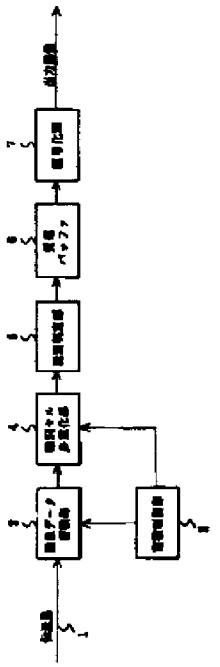
最終頁に続く

(54)【発明の名称】　画像蓄積伝送装置

(57)【要約】

【目的】　この発明は、画像データ蓄積部におけるセルのフレーム編集によりフレーム間のセルのシーケンス番号が非連続になった場合に、セルの廃棄と誤認するのを防ぐことを特徴とする。

【構成】　昇順にシーケンス番号が符された各セルを画像データ蓄積部２に蓄積した後、蓄積制御部３によって各セルをフレーム編集した際、前後して出力されるフレーム間のシーケンス番号が非連続になった時、識別セル多重化部４を稼働し、フレーム間に識別セルを多重化する。この結果、廃棄判定部７では識別セルを認識することで、シーケンス番号の非連続線にも拘らず、シーケンス番号欠落部分のセルの廃棄を誤認することがない。



1

【特許請求の範囲】

【請求項1】　個々に昇順のシーケンス番号が符されたセル単位の符号化画像データを伝送路を介して蓄積部へ蓄積した後、必要に応じてフレーム編集を施して復号化部へ出力する画像蓄積伝送装置において、上記蓄積部に蓄積画像データをフレーム編集する蓄積制御部と、フレーム編集によりシーケンス番号が非連続になったフレームの前に識別セルを多重する識別セル多重化部と、前記識別セル読み取りによりセルの廃棄を判定する廃棄判定部とを備えたことを特徴とする画像蓄積伝送装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】この発明は、伝送されてきた符号化画像データを一旦蓄積した後、復号化部へ伝送する画像蓄積伝送装置に関するものである。

【0002】

【従来の技術】以下に、本発明に関する従来技術の説明をする。図3は、テレビジョン学会誌、Ｖｏｌ．４３，Ｎｏ１０（１９８９）ｐｐ．１０７２～１０７８“周期的フレーム内符号化を用いた蓄積メディア用動画象符号化方式”および同Ｖｏｌ．４３，Ｎｏ１０（１９８９）ｐｐ．１１０６～１１１１“フレーム間符号化方式を用いた記録系動画再生システムの検討”の文中から予測した従来の画像蓄積再生装置の構成を示すものであるが、ここで3は蓄積媒体8の制御を行う蓄積制御部、9は蓄積媒体のクロックから復号クロックへ乗せ換えるための速度変換バッファ、7は符号化された画像データを可変長復号化し、さらに逆量子化を行い、モニタへ表示するための復号化部である。

【0003】図4は、Ｂ－ＩＳＤＮの伝送方式の一種であるＡＴＭ（Ａｓｙｎｃｈｒｏｎｏｕｓ　Ｔｒａｎｓｆｅｒ　Ｍｏｄｅ：非同期転送モード）網を用いた画像伝送装置の受信部の構成を表したものであり、1は伝送路（ＡＴＭ）、5は廃棄判定部、6は伝送路でのデータの揺らぎ等を吸収するための受信バッファ、7は符号化された画像データを可変長復号し、さらに逆量子化を行い、モニタへ表示するための復号化部、１０は音声多重化されたセルの分解、合成を行うセル分解・合成部である。図5は上記ＡＴＭ網での伝送概念の一例を表したものである。各セルは昇順にシーケンス番号が符下れて蓄積媒体8に蓄積された後、蓄積制御部3によって、所定数のセル単位毎にフレーム編集される。

【0004】次に従来技術の詳細動作について説明する。まず、従来の画像蓄積再生装置は図3のように構成されるが、ここで蓄積媒体8にはあらかじめフレーム間符号化された一定量のデータが蓄積されている。ここでは蓄積メディアとして例えば画像フレームの編集、コマ送り等ができるように、ここでの画像データは一定の周期でフレーム内符号化されている（例えば３０フレーム周期でフレーム内符号化をしていればフレーム内のみを

2

蓄積媒体から読み出せば、コマ送りとなる）。次に蓄積制御部3で蓄積媒体8からデータを読み出す制御を行う。読み出されたデータは速度変換バッファ9にてクロックの乗せ変えが行われ、復号部7で可変長復号化、逆量子化、データフォーマット変換等が行われ、モニタに出力される。

【0005】また、従来のＡＴＭ網を用いた画像伝送装置は図4のように構成されているが、ここでＡＴＭの伝送路1から入って来た画像画像セル化データはセル分解・合成部１０で音声／映像の分解網での制御情報などが分解される。次にＡＴＭは網が輻輳した場合、優先度の低いセルが廃棄されるため、廃棄判定部5にてセルの廃棄があったかどうかを判定する。これはセル1つ1つに対して付加されているシーケンス番号の連続性を見て判定する。廃棄判定部5から出力されたデータは受信バッファ6に格納される。ここでは伝送路1でのデータ揺らぎに対する許容、また伝送クロックから復号クロックに乗せ換える機能も司る。復号化部7では可変長符号化されて伝送されて来たデータを逆に可変長復号化し、逆量子化、データフォーマットの変換等を行い、モニタに出力する。

【0006】

【発明が解決しようとする課題】従来のＡＴＭ網用の画像伝送装置で、図3のように蓄積メディアを搭載した場合、以下のような問題点が生じる。即ち、蓄積媒体で蓄積された画像データをフレーム編集（コマ送り、逆再生等）して読み出した時、セルの順番が伝送路より入って来た時と異なってしまう。伝送路で廃棄が起こったかどうかは、先にも述べたが、セル1つ1つに対してヘッダとして付加されているシーケンス番号が連続しているかどうかで判定するので蓄積媒体でセルの順番を加工してしまうと、廃棄と誤認してしまうという問題があった。

【0007】この発明は上記のような問題点を解決するためになされたもので、蓄積部での編集で変わってしまったシーケンス番号の連続を認識し、廃棄との区別を行い、正しい廃棄処理の行える画像蓄積伝送装置を得ることを目的とする。

【0008】

【課題を解決するための手段】この発明に係る画像蓄積伝送装置は、個々に昇順のシーケンス番号が符されたセル単位の符号化画像データを伝送路を介して蓄積部へ蓄積した後、必要に応じてフレーム編集を施して復号化部へ出力する画像蓄積伝送装置において、上記蓄積部に蓄積画像データをフレーム編集する蓄積制御部と、フレーム編集によりシーケンス番号が非連続になったフレームの前に識別セルを多重する識別セル多重化部と、前記識別セル読み取りによりセルの廃棄を判定する廃棄判定部とを備えたものである。

【0009】

【作用】この発明における識別セル多重化部は、制御部

3

でフレーム編集シーケンス番号が非連続になったという情報（タイミング等）を受けてフレーム編集したフレームの前に識別セルを多重し、後段の廃棄判定部で編集かどうかを判断するものである。

【0010】

【実施例】以下、この発明の一実施例を図について説明する。図1において、1は網が輻輳した場合に、セルが廃棄される可能性のある伝送方式の網（例えばATM）（伝送路）、2は伝送路1より送られて来た画像データを一定量蓄積するための画像データ蓄積部、3は画像データ蓄積部2でのデータの書き込みと読み出しを制御する蓄積制御部、4は画像データ蓄積部2でフレーム編集等を行った場合に廃棄と区別するために識別セルを付加するための識別セル多重化部、5は伝送路でセルの廃棄があったかどうかを判定する廃棄判定部、6は伝送クロックから復号クロックへ乗せ換えるための受信バッファ、7は符号化された画像データを復号化しモニタへ表示するための復号化部である。

【0011】次に、この発明の一実施例について作用、動作を説明する。まず図1で、伝送路1から送られてきた画像データは、一旦画像データ蓄積部2で蓄積される。ここでの蓄積容量は画像フレームを複数フレーム蓄積できる容量を持っているものとする。この画像データ蓄積部2へデータを書き込む制御、および読み出す制御、また画像データ蓄積部2内でフレーム編集等を制御するのが蓄積制御部3であり、蓄積制御部3でフレーム編集した場合にそのフレームおよびタイミングを認識しているので、その情報を識別セル多重化部7に送る。画像データ蓄積部2から読み出されたデータ（セル）は識別セル多重化部4に入力される。ここでは、画像データ蓄積部2でフレーム編集を行わない場合、データはそのまま通過する。画像データ蓄積部2でフレーム編集を行った場合に各フレーム間のセルのシーケンス番号が非連続になった時は、蓄積制御部3で編集をしたタイミングで識別セルを多重化する。これは図2と図4に1例を表しており、それぞれの図について説明する。

【0012】まず図4であるが、これは通常のセルの並びであり、セル内の番号はパケットごと付いているシーケンス番号である。図4ではフレームごとに第1フレームから第3フレームまでを表示してあるが、ここでは編集を行って、第1フレームと第3フレームを接続した場合を述べる。図2は編集した場合のセルの並びである。第2フレームを削除してしまったために、第1フレームと第2フレームのシーケンス番号が連続でなくなっている。このため識別セルを多重しないと第2フレームは廃棄されたものとして扱われてしまう。識別セル多重化部7に手上記の処理を施した後、廃棄判定部4で廃棄されているかの判定を行い、速度変換バッファ5で伝送クロックから復号クロックに乗せ換える。そして復号部6で画像データの各種復号を行い、画像を表示するためのフォーマットに変換して出力する。

【0013】

【発明の効果】以上のように、この発明によれば蓄積部でのデータ編集を行った場合にセルのシーケンス番号が非連続になった時、フレーム編集を行ったフレームの前に識別セルを多重するように構成したので、セル廃棄または編集によるシーケンス番号の非連続を容易に区別でき、正しい廃棄処理が行われるという効果がある。

【図面の簡単な説明】

【図1】この発明の一実施例による画像蓄積伝送装置の構成図である。

【図2】この発明の一実施例によるセルフォーマットを示す図である。

【図3】従来の画像蓄積伝送装置の構成図である。

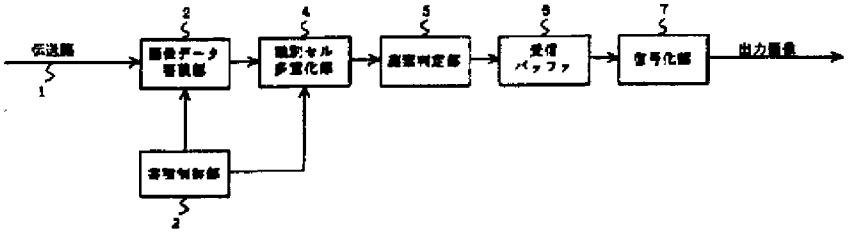
【図4】ATM用画像伝送装置の構成図である。

【図5】ATMのセルフォーマットを示す図である。

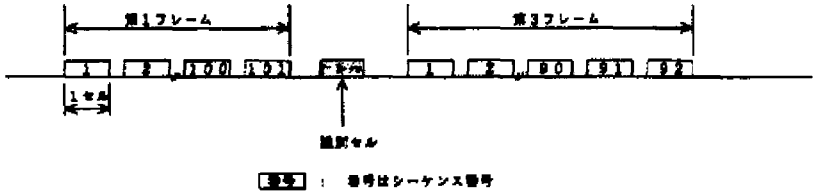
【符号の説明】

1　伝送路（ATM網）
2　画像データ蓄積部
3　蓄積制御部
4　識別セル多重化部
5　廃棄判定部
7　復号部
8　蓄積媒体
10　セル分解・合成部

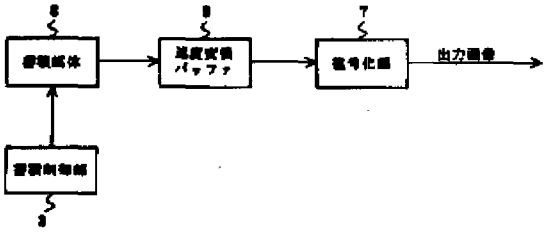
【図1】



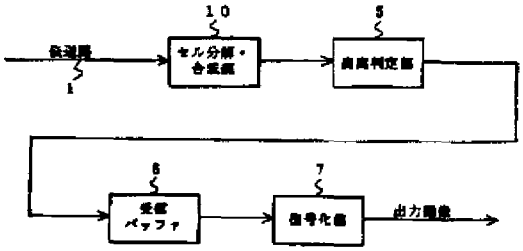
【図２】



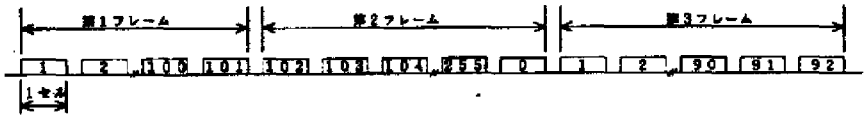
【図３】



【図４】



【図５】



【手続補正書】
【提出日】平成３年１２月１０日
【手続補正１】
【補正対象書類名】明細書
【補正対象項目名】０００３
【補正方法】変更
【補正内容】
【０００３】図４は、Ｂ－ＩＳＤＮの伝送方式の一種であるＡＴＭ（Ａｓｙｎｃｈｒｏｎｏｕｓ　Ｔｒａｎｓｆｅｒ　Ｍｏｄｅ：非同期転送モード）網を用いた画像伝送装置の受信部の構成を表したものであり、１は伝送路（ＡＴＭ）、５は廃棄判定部、６は伝送路でのデータの揺らぎ等を吸収するための受信バッファ、７は符号化された画像データを可変長復号し、さらに逆量子化を行い、モニタへ表示するための復号化部、１０は音声多重化されたセルの分解、合成を行うセル分解・合成部である。図５は上記ＡＴＭ網での伝送概念の一例を表したものである。各セルは昇順にシーケンス番号が付加されて蓄積媒体８に蓄積された後、蓄積制御部３によって、所定数のセル単位毎にフレーム編集される。

フロントページの続き

(72)発明者　猪股　英樹
神奈川県鎌倉市大船五丁目１番１号　三菱
電機株式会社通信システム研究所内

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant:	HANNUKSELA	Examiner:	Lee, S.
Serial No.:	11/338,934	Group Art Unit:	2611
Filed:	January 25, 2006	Docket No.:	TPO.004.C2
Confirmation No.:	6610	Customer No.:	76385
Title:	GROUPING OF IMAGE FRAMES IN VIDEO CODING		

SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT
UNDER 37 C.F.R. §1.97(c)(1) and 37 C.F.R. §1.97 (e)(1)

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

With regard to the above-identified application, the items of information listed on the enclosed Form 1449 are brought to the attention of the Examiner.

This statement should be considered because it is submitted after the mailing date of the first Office Action on the merits but before the mailing date of a final Office Action under 37 C.F.R. §1.113 or a Notice of Allowance under 37 C.F.R. §1.311.

The items listed on the enclosed Form 1449 were first cited in a communication from a foreign patent office regarding a counterpart foreign application not more than three months prior to the filing of the information disclosure statement. Accordingly, no fee is due for consideration of the items listed on the enclosed Form 1449. However, if fees are deemed necessary, authorization is given to charge Deposit Account 50-3581 (TPO.004.C2) for consideration of the information disclosure statement.

In accordance with 37 C.F.R. § 1.98(a)(2), only a copy of each foreign document or non-U.S. patent/application listed on the enclosed Form 1449 is provided.

Please note that any notations or markings on the attached documents do not reflect particular relevance, or lack thereof, to the present application, nor were they necessarily made by anyone affiliated with the prosecution of the present application.

No representation is made that a reference is "prior art" within the meaning of 35 U.S.C. §§ 102 and 103. In addition, Applicant does not represent that a reference has been thoroughly reviewed or that any relevance of any portion of a reference is intended, and reserves the right to establish otherwise under 37 C.F.R. § 1.131 or others.

Consideration of the items listed is respectfully requested. According to MPEP § 609, Applicant requests that the Examiner return a copy of the attached Form 1449, marked as being considered and initialed by the Examiner, to the undersigned with the next official communication.

Respectfully submitted,

HOLLINGSWORTH & FUNK, LLC
8500 Normandale Lake Blvd., Suite 320
Minneapolis, MN 55437
952.854.2700

By: /Steven R. Funk/
Steven R. Funk
Reg. No. 37,830

Electronic Acknowledgement Receipt

EFS ID:	8624884
Application Number:	11338934
International Application Number:	
Confirmation Number:	6610
Title of Invention:	Grouping of image frames in video coding
First Named Inventor/Applicant Name:	Miska Hannuksela
Customer Number:	76385
Filer:	Steven R. Funk
Filer Authorized By:	
Attorney Docket Number:	TPO.004.C2
Receipt Date:	14-OCT-2010
Filing Date:	25-JAN-2006
Time Stamp:	14:33:35
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Information Disclosure Statement (IDS) Filed (SB/08)	Form1449-12TPO004C2.pdf	71878 07c1f412c9687312dbfb7b4a384b8258d601a372	no	1

Warnings:	
Information:	

This is not an USPTO supplied IDS fillable form					
2	Foreign Reference	JP11275576pre.pdf	236628 f872f28d2862519b5f1d6ad930ba7863892ec8e3	no	35
Warnings:					
Information:					
3	Foreign Reference	JP4348690pre.pdf	188895 cabd1429d49abfe3242ac88d82326510350983ea	no	6
Warnings:					
Information:					
4	NPL Documents	JPOA20100817.pdf	74168 f318ac7a8fbcd4fe55a3d0fd18432cfe25dd0936	no	4
Warnings:					
Information:					
5	NPL Documents	JPOA20100824.pdf	94278 a850e4ac8e953ae55e3060329b123b4f4e6e0fda	no	5
Warnings:					
Information:					
6	Transmittal Letter	IDS-12-TPO004C2.pdf	57754 e6ab8a976ba1a5f7718b6cdd231ab7eeab1a13a4	no	2
Warnings:					
Information:					
		Total Files Size (in bytes):	723601		

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111

If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371

If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office

If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
11/338,934	01/25/2006	Miska Hannuksela	TPO.004.C2	6610

10888 7590 12/03/2010

Hollingsworth & Funk
8500 Normandale Lake Blvd., Suite 320
Minneapolis, MN 55437

EXAMINER	
LEE, SIU M	

ART UNIT	PAPER NUMBER
2611	

MAIL DATE	DELIVERY MODE
12/03/2010	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary	Application No.	Applicant(s)	
	11/338,934	HANNUKSELA, MISKA	
	Examiner	Art Unit	
	SIU M. LEE	2611	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE *3* MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *17 September 2010*.

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1,2,4-6,8-11,13 and 14* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *5,6,8 and 13* is/are rejected.

7)☒ Claim(s) *1,2,4,9-11 and 14* is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *25 January 2006* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☒ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☒ Certified copies of the priority documents have been received in Application No. 10/306,942.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☐ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date *See Continuation Sheet*.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

Continuation of Attachment(s) 3). Information Disclosure Statement(s) (PTO/SB/08), Paper No(s)/Mail Date :8/10/2010, 9/30/2010, 10/14/2010.

DETAILED ACTION

Response to Arguments

1. Applicant's arguments with respect to the 35 U.S.C. 101 rejection of claims 5-8 and 13, filed 9/17/2010 have been fully considered but they are not persuasive.

Applicant's argument:

The claimed encoder and decoder are well known in the art to refer to a device or circuit; the specification illustrates structural component of the encoder (e.g. compressor 108 and multiplexer 110) and decoder (e.g. demultiplexer 116 and decompressor 118). Paragraph 0095 teaches example of a decoder including a low bit rate video decoder as described in the ITU-T recommendation H.263 or H.26L. Moreover, paragraph 0096 describes positioning of the encoder and decoder that would not be possible if they were "pure software".

Examiner's response:

The following recites paragraph 0104 of the publication of the instant application (US 2006/0120451 A1):

> [0104] It is to be noted that according to the invention the functional elements of the above described streaming system and its elements, such as the streaming server, video encoder, video decoder and terminal are preferably implemented by means of software, by hardware solutions, or as a combination of the two. The coding and decoding methods of the invention are particularly suitable for implementation as computer software comprising computer-readable commands for executing the process steps of the invention. A preferred way of implementing the encoder and the decoder is to store them in a storage means as a program code that can be executed by a computer-like device, for example a personal computer (PC) or a mobile station, to provide coding/decoding functionalities on the device in question.

Paragraph 0104 discloses that the video encoder, video decoder and terminal are preferably implemented by means of software, by hardware solutions, or as a combination of the two. Since claim 5 is non-structural per se and is only limited by the functions, therefore, a reasonable broadest interpretation in light of the specification leads to the conclusion that the claimed video encoder covers a hardware embodiment (statutory) as well as a software embodiment (non-statutory), which does not fall within the definition of a process, machine, manufacture, or composition of matter.

With regard to the argument that the specification illustrates structural component of the encoder (e.g. compressor 108 and multiplexer 110), the components are represented by block in figure 1 and only define by the function as disclose in the specification, therefore, it does not exclude a software implementation of these components.

With respect to the argument that Paragraph 0095 teaches example of a decoder including a low bit rate video decoder as described in the ITU-T recommendation H.263

or H.26L. Moreover, paragraph 0096 describes positioning of the encoder and decoder that would not be possible if they were "pure software"; however, it is discloses in the above quoted paragraph (paragraph 0104 of the publication of the instant application (US 2006/0120451 A1)) that encoder and decoder can be implemented as software, as the claim is not structural per se, the examiner interprets the claim covers a software embodiment as well, therefore the rejection is maintained. If the applicant considered that the encoder and decoder can not be “pure software’, the examiner suggests amending paragraph 0104 of the publication of the instant application to remove the software embodiment.

2. Applicant's arguments with respect to the 35 U.S.C. 112, 1st paragraph rejection of claims 5-6, 8, and 13 filed 9/17/2010 have been fully considered but they are not persuasive.

Applicant argument:

Applicant respectfully traverses the § 112, first paragraph, rejection of claims 5, 6, 8, and 13 because the claims are not "single means" claims. First, none of these claims uses a "means recitation" or "means plus function" language. While a claim dependent upon a recited property could result in a fact situation similar to that discussed in the cited Hyatt decision when the claim covers every conceivable structure for achieving the stated property, such is not the case for claims 5, 6, 8, and 13. In contrast the claims are limited to the structures of an encoder (claims 5, 6, and 8) and of a decoder (claim 13), which are well-accepted and understood terms in the art, for

achieving the stated properties of the respective claims. Therefore, the claims are not unduly broad and are enabled by the specification. Applicant accordingly requests that the rejection be withdrawn.

Examiner's response:

The examiner would like to point out that claim 5-6, 8, and 13 are not rejection as a single means claim rejection. Claims 5-6, 8, and 13 are rejected based on the claim only directed to a single structure or component (decoder) and the decoder is defined by the recited properties (functions of the decoder), than a fact situation similar to that discussed in the cited Hyatt decision when the claim covers every conceivable structure for achieving the stated property for achieving the stated property while the specification discloses at most only known to the invention. Therefore, the 35 U.S.C. 112, 1st paragraph rejections of claims 5-6, 8, and 13 are maintained.

3. Applicant's arguments, see page 7-9, filed 9/17/2010, with respect to 35 U.S.C. 102(e) and 103(a) rejections of claims 1-16 have been fully considered and are persuasive. The 35 U.S.C. 102(e) and 103(a) rejections of claims 1-16 has been withdrawn.

With regard to the double patenting rejection of claims 1-2 and 4 are withdrawn because of the combination of claim 3 into claim 1.

Claim Objections

4. Claims 1-2, 4-6, 8, 10-11, are objected to because of the following informalities:

Claim 1, line 9 recites "encoding identifier values for image frames"; since "image frames" was introduced in line 2, the examiner suggests changing to "encoding identifier values for **the** image frames".

Claim 1, line 10-11 recite "resetting the identifier value for the indicated first image frame of an independent sequence"; since "independent sequence" was introduced in line 2, it is unclear whether it is the same independent sequence or another independent sequence.

Claim 5, line 9 recites "encoding identifier values for image frames"; since "image frames" was introduced in line 2, the examiner suggests changing to "encoding identifier values for **the** image frames".

Claim 5, line 10-11 recite "resetting the identifier value for the indicated first image frame of an independent sequence"; since "independent sequence" was introduced in line 2, it is unclear whether it is the same independent sequence or another independent sequence.

Claim 9, line 10 recites "encoding identifier values for image frames"; since "image frames" was introduced in line 3, the examiner suggests changing to "encoding identifier values for **the** image frames".

Claim 9, line 12-13 recite "resetting the identifier value for the indicated first image frame of an independent sequence"; since "independent sequence" was introduced in line 2, it is unclear whether it is the same independent sequence or another independent sequence.

Claim 10, lines 12-13 recite "resetting the identifier value for the indicated first image frame of an independent sequence"; since "independent sequence" was introduced in line 6, it is unclear whether it is the same independent sequence or another independent sequence.

Claim 13, lines 13-14 recite "resetting the identifier value for the indicated first image frame of an independent sequence"; since "independent sequence" was introduced in line 7, it is unclear whether it is the same independent sequence or another independent sequence.

Claim 14, lines 15-16 recite "resetting the identifier value for the indicated first image frame of an independent sequence"; since "independent sequence" was introduced in lines 8-9, it is unclear whether it is the same independent sequence or another independent sequence.

Appropriate correction is required.

Claim Rejections - 35 USC § 101

5. 35 U.S.C. 101 reads as follows:

> Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

Claims 5-8 are rejected under 35 U.S.C. 101 because the claimed invention is directed to non-statutory subject matter. The claim is NOT to a process, machine, manufacture or composition of matter.

The following recites paragraph 0104 of the publication of the instant application (US 2006/0120451 A1):

> [0104] It is to be noted that according to the invention the functional elements of the above described streaming system and its elements, such as the streaming server, video encoder, video decoder and terminal are preferably implemented by means of software, by hardware solutions, or as a combination of the two. The coding and decoding methods of the invention are particularly suitable for implementation as computer software comprising computer-readable commands for executing the process steps of the invention. A preferred way of implementing the encoder and the decoder is to store them in a storage means as a program code that can be executed by a computer-like device, for example a personal computer (PC) or a mobile station, to provide coding/decoding functionalities on the device in question.

Paragraph 0104 discloses that the video encoder, video decoder and terminal are preferably implemented by means of software, by hardware solutions, or as a combination of the two. Since the encoder of claim 5 is non-structural per se and is only limited by the functions, therefore, a reasonable broadest interpretation in light of the specification leads to the conclusion that the claimed video encoder covers a hardware embodiment (statutory) as well as a software embodiment (non-statutory), which does not fall within the definition of a process, machine, manufacture, or composition of matter.

Claim 13 is rejected under 35 U.S.C. 101 because the claimed invention is directed to non-statutory subject matter. The claim is NOT to a process, machine, manufacture or composition of matter.

The following recites paragraph 0104 of the publication of the instant application (US 2006/0120451 A1):

> [0104] It is to be noted that according to the invention the functional elements of the above described streaming system and its elements, such as the streaming server, video encoder, video decoder and terminal are preferably implemented by means of software, by hardware solutions, or as a combination of the two. The coding and decoding methods of the invention are particularly suitable for implementation as computer software comprising computer-readable commands for executing the process steps of the invention. A preferred way of implementing the encoder and the decoder is to store them in a storage means as a program code that can be executed by a computer-like device, for example a personal computer (PC) or a mobile station, to provide coding/decoding functionalities on the device in question.

Paragraph 0104 discloses that the video encoder, video decoder and terminal are preferably implemented by means of software, by hardware solutions, or as a combination of the two. Since the decoder of claim 13 is non-structural per se and is only limited by the functions, therefore, a reasonable broadest interpretation in light of the specification leads to the conclusion that the claimed video decoder covers a hardware embodiment (statutory) as well as a software embodiment (non-statutory), which does not fall within the definition of a process, machine, manufacture, or composition of matter.

Claim Rejections - 35 USC § 112

6. The following is a quotation of the first paragraph of 35 U.S.C. 112:

> The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.

Claims 5-8 are rejected under 112, 1st paragraph as being a single component (video encoder) claim, i.e., where a single component (video encoder) recitation does not appear in combination with another recited component, is subject to an undue breadth rejection. The single component which covered every conceivable ways for achieving the stated purpose was held nonenabling for the scope of claim because the specification disclosed at most only those known to the inventor. See MPEP §2164.08(a).

Claim 13 is rejected under 112, 1st paragraph as being a single component (video decoder) claim, i.e., where a single component (video decoder) recitation does not appear in combination with another recited component, is subject to an undue breadth rejection. The single component which covered every conceivable ways for achieving the stated purpose was held nonenabling for the scope of claim because the specification disclosed at most only those known to the inventor. See MPEP §2164.08(a).

Allowable Subject Matter

7. Claims 1-2, 4, 9-11, and 14 are objected to because of minor informality, but would be allowable if rewritten to overcome the objection.

8. The following is a statement of reasons for the indication of allowable subject matter:

The present invention describes a method for encoding a video sequence comprising an independent sequence of image frames, wherein at least one reference

image frame is predictable from at least one previous image frame, and wherein the at least one reference image frame is earlier than the at least one previous image frame in decoding order, the method comprising encoding into the video sequence an indication of at least one image frame, which is the first image frame, in decoding order, of the independent sequence, said at least one reference image frame being included in the sequence; encoding identifier values for image frames according to a numbering scheme; and resetting the identifier value for the indicated first image frame of an independent sequence. the closest prior art, Kim in view of Ueda and Kawashima together disclose a similar method but fails to disclose encoding identifier values for image frames according to a numbering scheme; and resetting the identifier value for the indicated first image frame of an independent sequence.

The present invention describes a method for decoding a compressed video sequence, wherein at least one reference image frame is predictable from at least one previous image frame and wherein the at least one reference frame is earlier than the at least one previous image frame in decoding order, the method comprising decoding from the video sequence an indication of at least one image frame, which is the first image frame, in decoding order, of an independent sequence, said sequence including the at least one reference image frame; starting the decoding of the video sequence from said first image frame of the independent sequence, whereby the video sequence is decoded without prediction from any image frame decoded prior to said first image frame; decoding identifier values for image frames according to a numbering scheme; and resetting the identifier value for the indicated first image frame of an independent

sequence. The closest prior art, Kim in view of Ueda and Kawashima together disclose a similar method but fails to disclose decoding identifier values for image frames according to a numbering scheme; and resetting the identifier value for the indicated first image frame of an independent sequence.

Conclusion

9. **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to SIU M. LEE whose telephone number is (571)270-1083. The examiner can normally be reached on Mon-Fri, 7:30-4:00 with every other Friday off.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Chieh Fan can be reached on (571) 272-3042. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Siu M Lee/
Examiner, Art Unit 2611
12/1/2010

/CHIEH M FAN/

Supervisory Patent Examiner, Art Unit 2611

Index of Claims	Application/Control No.	Applicant(s)/Patent Under Reexamination
	11338934	HANNUKSELA, MISKA
	Examiner	Art Unit
	SIU M LEE	2611

✓	Rejected	-	Cancelled	N	Non-Elected	A	Appeal
=	Allowed	÷	Restricted	I	Interference	O	Objected

☐ Claims renumbered in the same order as presented by applicant ☐ CPA ☐ T.D. ☐ R.1.47

CLAIM		DATE								
Final	Original	05/21/2010	12/02/2010							
	1	✓	O							
	2	✓	O							
	3	✓	-							
	4	✓	O							
	5	✓	✓							
	6	✓	✓							
	7	✓	-							
	8	✓	✓							
	9	✓	O							
	10	✓	O							
	11	✓	O							
	12	✓	-							
	13	✓	✓							
	14	✓	O							
	15	✓	-							
	16	✓	-							

Search Notes	Application/Control No.	Applicant(s)/Patent Under Reexamination
	11338934	HANNUKSELA, MISKA
	Examiner	Art Unit
	SIU M LEE	2611

SEARCHED

Class	Subclass	Date	Examiner
375	240.12, 240.25	5/21/2010	Siu M. Lee

SEARCH NOTES

Search Notes	Date	Examiner
EAST	5/21/2010	Siu M. Lee
Inventor name search - Miska hannuksela	5/21/2010	Siu M. Lee
Discuss the application with Chieh Fan	5/21/2010	Siu M. Lee
EAST	12/1/2010	Siu M. Lee
Discuss the amendment with Chieh Fan	12/1/2010	Siu M. Lee
Discuss the 101 and 112 rejection of claims 5-8 and 13 with John Peng	12/1/2010	Siu M. Lee

INTERFERENCE SEARCH

Class	Subclass	Date	Examiner

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

BIB DATA SHEET

CONFIRMATION NO. 6610

SERIAL NUMBER	FILING or 371(c) DATE RULE	CLASS	GROUP ART UNIT	ATTORNEY DOCKET NO.
11/338,934	01/25/2006	375	2611	TPO.004.C2

APPLICANTS
Miska Hannuksela, Tampere, FINLAND;

**** CONTINUING DATA ***************************
This application is a CON of 10/306,942 11/29/2002

**** FOREIGN APPLICATIONS ***************************
FINLAND 20020127 01/23/2002

**** IF REQUIRED, FOREIGN FILING LICENSE GRANTED ****
02/28/2006

Foreign Priority claimed ☑ Yes ☐ No 35 USC 119(a-d) conditions met ☑ Yes ☐ No Verified and Acknowledged /SIU M LEE/ Examiner's Signature	☐ Met after Allowance Initials	STATE OR COUNTRY	SHEETS DRAWINGS	TOTAL CLAIMS	INDEPENDENT CLAIMS
		FINLAND	4	16	7

ADDRESS
Hollingsworth & Funk
8500 Normandale Lake Blvd., Suite 320
Minneapolis, MN 55437
UNITED STATES

TITLE
Grouping of image frames in video coding

FILING FEE RECEIVED		
1800	FEES: Authority has been given in Paper No._____ to charge/credit DEPOSIT ACCOUNT No._____ for following:	☐ All Fees ☐ 1.16 Fees (Filing) ☐ 1.17 Fees (Processing Ext. of time) ☐ 1.18 Fees (Issue) ☐ Other _____ ☐ Credit

FORM 1449* INFORMATION DISCLOSURE STATEMENT IN AN APPLICATION (Use several sheets if necessary)	Docket Number: TPO.004.C2	Application Number: 11/338,934
	Applicant: HANNUKSELA	
	Filing Date: 1/25/2006	Group Art Unit: 2611

U.S. PATENT DOCUMENTS

EXAMINER INITIAL	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
/SL/	5852630	12/22/1998	Langberg			
/SL/	6072831	06/06/2000	Chen			
/SL/	6363208	03/26/2002	Nitta et al.			

FOREIGN PATENT DOCUMENTS

	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	TRANSLATION	
						YES	NO
/SL/	2137197C1	9/10/1999	Russia				

OTHER DOCUMENTS (Including Author, Title, Date, Pertinent Pages, Etc.)

/SL/	1997	Illgner et al., "Spatially Scalable Video Compression Employing Resolution Pyramids", IEEE Journal on Selected Areas in Communications, Vol. 15, No. 9, December 1, 2007, pg. 1688-1703.
/SL/	2000	Information Technology – Generic Coding of Moving Pictures and Associated Audio Information: Systems, ISO/IEC 13818-1, Second edition, December 1, 2000, 174 pages.
/SL/	5/5/2010	Notice of Allowance dated 5/5/2010 from Russian Application No. 2006110321/09, 11 pages.
/SL/	6/17/2010	Office Action dated 6/17/2010 from U.S. Serial No. 11/338,996, 30 pages.
/SL/	6/23/2010	Office Action dated 6/23/2010 from U.S. Serial No. 10/306,942, 6 pages.
/SL/	5/25/2010	Office Action dated 5/25/2010 from U.S. Serial No. 10/306,942, 6 pages.
/SL/	4/09/2009	Office Action dated 4/9/2009 from U.S. Serial No. 10/306,942, 4 pages.
/SL/	5/8/2006	Office Action Response dated 5/8/2006 from U.S. Serial No. 10/306,942, 11 pages.

EXAMINER /Siu Lee/	DATE CONSIDERED 12/01/2010

EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609; draw line through citation if not in conformance and not considered. Include copy of this form for next communication to the Applicant.

FORM 1449* SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT IN AN APPLICATION (Use several sheets if necessary)	Docket Number: TPO.004.C2	Application Number: 11/338,934
	Applicant: HANNUKSELA	
	Filing Date: 01/25/2006	Group Art Unit: 2611

U.S. PATENT DOCUMENTS

EXAMINER INITIAL	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE

FOREIGN PATENT DOCUMENTS

	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	TRANSLATION	
						YES	NO
/SL/	JP11275576	10/08/1999	JP			Abstract only	
/SL/	JP4348690	12/03/1992	JP			Abstract only	

OTHER DOCUMENTS (Including Author, Title, Date, Pertinent Pages, Etc.)

/SL/	08/17/2010	Copy of translated Office Action dated 08/17/2010 from parallel Japanese Application No. 2006-120297, 4 pages.
/SL/	08/24/2010	Copy of translated Office Action dated 08/24/2010 from parallel Japanese Application No. 2006-120296, 5 pages.

EXAMINER /Siu Lee/	DATE CONSIDERED 12/01/2010

EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609; draw line through citation if not in conformance and not considered. Include copy of this form for next communication to the Applicant.

*Substitute Disclosure Statement Form (PTO-1449) Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

EAST Search History

EAST Search History (Prior Art)

Ref #	Hits	Search Query	DBs	Default Operator	Plurals	Time Stamp
S1	1	11/338934 and program	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/19 14:20
S2	190	(miska near2 Hannuksela).in.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/19 15:30
S3	149	S2 and stream$3	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/19 15:30
S4	38	S3 and (@pd<="20020123" OR @ad<="20020123" OR @rlad<="20020123" OR @prad<="20020123")	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/19 15:31
S5	2	"7289506".pn.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/19 15:33
S6	1	09/560217	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/19 15:33
S7	21030	encod$3 with (picture or frame) with (indicat$3 or tag or flag or reference)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 09:17

S8	12392	S7 and (@pd<="20020123" OR @ad<="20020123" OR @rlad<="20020123" OR @prad<="20020123")	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 09:17
S9	567	S8 and (encod$3 with (picture or frame) with ((indicat$3 or tag or flag or reference) near3 first))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 09:18
S10	43406	scalability	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 09:19
S11	45	S9 and S10	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 09:19
S12	2	"7236526".pn.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 09:21
S13	7243	encod$3 with (picture or frame) with (indicator or tag or flag or label$4)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 09:41
S14	4794	S13 and (@pd<="20020123" OR @ad<="20020123" OR @rlad<="20020123" OR @prad<="20020123")	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 09:41
S15	149	S14 and scalability	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 09:41
S16	12	S15 and ((indicat$3 or tag or flag or reference) near3 (I or independent or first) frame)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 09:42

S30	275	(indicator or flag or identifier or tag or bit) with header with (i or independent or first) frame	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 11:52
S31	136	S30 and (@pd<="20020123" OR @ad<="20020123" OR @rlad<="20020123" OR @prad<="20020123")	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 11:52
S32	331	(indicator or flag or identifier or tag or bit) with header with (i or independent or first or intra) frame	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 11:52
S33	177	S32 and (@pd<="20020123" OR @ad<="20020123" OR @rlad<="20020123" OR @prad<="20020123")	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 11:52
S34	128	S33 and video	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 11:53
S35	2	"7336787".pn.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 12:07
S36	1	11/338934 and (medium)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 13:31
S37	1	11/338934 and (computer readable medium)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 13:31
S38	751	identifier with flag with header	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 14:47

S39	1	identifier with flag with header with video sequence	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 14:47
S40	4	identifier with flag with header with start with video	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 14:48
S41	20	encod$3 with identifier with flag with header	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 14:49
S42	2	resetting with (identifier or tag or indicator or flag) with independent with sequence	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 15:09
S43	9	video sequence with (identifier or tag or indicator or flag) with independent with sequence	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 15:36
S44	2	"6968005".pn.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/20 15:39
S50	4706	375/240.12,240.25.ccls.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/21 07:33
S51	2292	S50 and (@pd<="20020123" OR @ad<="20020123" OR @rlad<="20020123" OR @prad<="20020123")	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/21 07:33
S52	39	S51 and (video near2 encod$3) and (first frame with (indicat$3 or tag or identifier or flag))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/21 07:35

S53	2	"6094456".pn. and identifier	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/21 09:19
S54	12	10/092340 and decod$3	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/05/21 10:11
S55	1	11/338934 and (software or program)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/12/01 09:00
S56	1	11/338934 and reset$4	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/12/01 09:07
S57	0	"20030063806".pn. and reset$4	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/12/01 09:10
S58	25	(encod$3 with number with image with frame) same (reset $4)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/12/01 12:31
S59	0	reset$4 with frame number$3 with new sequence	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/12/01 12:34
S60	6	reset$4 with frame with number $3 with new sequence	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/12/01 12:34
S61	3	(encod$3 with video frame with number) same2 (reset$3 with (new or another) with (sequence or video))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/12/01 12:36

S62	82	reset$3 with (video or image) frame with number	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/12/01 12:41
S63	32	S62 and (@pd<="20020123" OR @ad<="20020123" OR @rlad<="20020123" OR @prad<="20020123")	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/12/01 12:41
S64	93	initializ$5 with ((image or video) near2 frame) with new	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/12/01 12:58
S65	1	initializ$5 with ((image or video) near2 frame) with (new near2 sequence)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/12/01 12:58
S66	47	S64 and (@pd<="20020123" OR @ad<="20020123" OR @rlad<="20020123" OR @prad<="20020123")	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/12/01 12:58
S67	5	(reset$4 or renew$3 or initializ $5 or re?initializ$5) with (tag or identit$3) with (image or video) frame	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/12/01 13:08
S68	1	11/338934	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/12/01 13:33

EAST Search History (Interference)

Ref #	Hits	Search Query	DBs	Default Operator	Plurals	Time Stamp
S45	140	(miska near2 Hannuksela).in.	US-PGPUB; USPAT; UPAD	ADJ	ON	2010/05/20 06:40
S46	139	S45 and stream$3	US-PGPUB; USPAT; UPAD	ADJ	ON	2010/05/20 06:41

S47	37	S46 and (@pd<="20020123" OR @ad<="20020123" OR @rlad<="20020123" OR @prad<="20020123")	US-PGPUB; USPAT; UPAD	ADJ	ON	2010/05/20 06:41
S48	37	S47 not "s4"	US-PGPUB; USPAT; UPAD	ADJ	ON	2010/05/20 06:41
S49	18599	encod$3 with (picture or frame) with (indicat$3 or tag or flag or reference)	US-PGPUB; USPAT; UPAD	ADJ	ON	2010/05/20 09:17

12/2/2010 7:09:11 AM
C:\Documents and Settings\slee7\My Documents\EAST\Workspaces\11338934 201005.wsp

FORM 1449* INFORMATION DISCLOSURE STATEMENT IN AN APPLICATION (Use several sheets if necessary)	Docket Number: TPO.004.C2	Application Number: 11/338,934
	Applicant: HANNUKSELA	
	Filing Date: 1/25/2006	Group Art Unit: 2611

U.S. PATENT DOCUMENTS

EXAMINER INITIAL	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE

FOREIGN PATENT DOCUMENTS

	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	TRANSLATION	
						YES	NO

OTHER DOCUMENTS (Including Author, Title, Date, Pertinent Pages, Etc.)

/SL/	9/15/2010	Office Action Response dated 9/15/2010 from U.S. Serial No. 10/306,942, 12 pages.
/SL/	9/10/2010	Office Action Response dated 9/10/2010 from U.S. Serial No. 11/338,996, 10 pages.

EXAMINER /Siu Lee/	DATE CONSIDERED 12/01/2010

EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609; draw line through citation if not in conformance and not considered. Include copy of this form for next communication to the Applicant.

*Substitute Disclosure Statement Form (PTO-1449) Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

IN THE UNITED STATES PATENT & TRADEMARK OFFICE

Applicant:	HANNUKSELA	Examiner:	Lee, S.
Serial No.:	11/338,934	Group Art Unit:	2611
Filed:	January 25, 2006	Docket No.:	TPO.004.C2
Confirmation No.:	6610	Customer No.:	10888
Title:	GROUPING OF IMAGE FRAMES IN VIDEO CODING		

CERTIFICATE UNDER 37 CFR 1.8: The undersigned hereby certifies that this paper is being electronically transmitted by EFS-WEB to the United States Patent and Trademark Office on February 2, 2011.

By: /Erin Nichols Matkaiti/
Erin Nichols Matkaiti

AMENDMENT & RESPONSE UNDER 37 C.F.R. § 1.116

Mail Stop AF
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

In response to the Office action mailed December 3, 2010, setting a three (3) month shortened statutory period for response, Applicant submits the present amendment and response under 37 C.F.R. § 1.116.

Claims 1, 2, 4-6, 8-11, 13, and 14 remain pending in the application. Claims 1, 5, 9, 10, 13, and 14 have been amended. No claims have been added or canceled. Applicant respectfully requests reconsideration and allowance of all pending claims.

AMENDMENTS TO THE SPECIFICATION begin on page 2.

AMENDMENTS TO THE CLAIMS are reflected in the listing of claims which begins on page 3 of this paper.

REMARKS/ARGUMENTS begin on page 8.

AMENDMENTS TO THE SPECIFICATION

Please amend the paragraph at page 33, lines 6-16, of the original specification as indicated below.

[0097] It is to be noted that according to the invention the functional elements of the above described streaming system and its elements, such as the streaming server, video encoder, video decoder and terminal are preferably implemented by ~~means of software, by~~ hardware solutions[[,]] or as a combination of ~~the two~~hardware and software. The coding and decoding methods of the invention are particularly suitable for implementation as computer software comprising computer-readable commands for executing the process steps of the invention. A preferred way of implementing the encoder and the decoder is to store them in a storage means as a program code that can be executed by a computer-like device, for example a personal computer (PC) or a mobile station, to provide coding/decoding functionalities on the device in question.

AMENDMENTS TO THE CLAIMS

Applicant respectfully requests that the amendments presented below be entered and that this listing of claims replace all prior versions of claims in the application.

1. (Currently amended) A method for encoding a video sequence comprising an independent sequence of image frames, wherein at least one reference image frame is predictable from at least one previous image frame, and wherein the at least one reference image frame is earlier than the at least one previous image frame in decoding order, the method comprising:

encoding into the video sequence an indication of at least one image frame, which is the first image frame, in decoding order, of the independent sequence, said at least one reference image frame being included in the sequence;

encoding identifier values for the image frames according to a numbering scheme; and

resetting the identifier value for the indicated first image frame of [[an]]the independent sequence.

2. (Original) A method according to claim 1, further comprising:

encoding the indication into the video sequence as a separate flag included in the header of a slice.

3. (Canceled)

4. (Original) A method according to claim 1, further comprising:

encoding into the video sequence an identifier value for the independent sequence.

5. (Currently amended) A video encoder ~~for encoding~~comprising at least one processor and at least one memory including computer program code, the at least one memory and the computer program code configured to, with the at least one processor, cause the video

encoder to encode a video sequence comprising an independent sequence of image frames, wherein at least one reference image frame is predictable from at least one previous image frame, and wherein the at least one reference frame is earlier than the at least one previous image frame in decoding order~~, the video encoder being configured to~~by causing the video encoder to at least:

encode into the video sequence an indication of at least one image frame, which is the first image frame, in decoding order, of the independent sequence, said at least one reference image frame being included in the sequence;

encode identifier values for the image frames according to a numbering scheme; and

reset the identifier value for the indicated first image frame of [[an]]the independent sequence.

6. (Previously presented) A video encoder according to claim 5, wherein the video encoder is configured to

encode the indication into the video sequence as a separate flag included in the header of a slice.

7. (Canceled)

8. (Previously presented) A video encoder according to claim 5, wherein the video encoder is configured to

encode into the video sequence an identifier value for the independent sequence.

9. (Currently amended) A computer program product, stored on a non-transitory computer readable medium and executable in a data processing device, for encoding a video sequence comprising an independent sequence of image frames, wherein at least one reference image frame is predictable from at least one previous image frame and wherein the at least one reference frame is earlier than the at least one previous image frame in decoding order, the computer program product comprising:

a computer program code for encoding into the video sequence an indication of at least one image frame, which is the first image frame, in decoding order, of the independent sequence, said at least one reference image frame being included in the sequence;

a computer code for encoding identifier values for the image frames according to a numbering scheme; and

a computer code for resetting the identifier value for the indicated first image frame of [[an]]the independent sequence.

10. (Currently amended) A method for decoding a compressed video sequence, wherein at least one reference image frame is predictable from at least one previous image frame and wherein the at least one reference frame is earlier than the at least one previous image frame in decoding order, the method comprising:

decoding from the video sequence an indication of at least one image frame, which is the first image frame, in decoding order, of an independent sequence, said sequence including the at least one reference image frame;

starting the decoding of the video sequence from said first image frame of the independent sequence, whereby the video sequence is decoded without prediction from any image frame decoded prior to said first image frame;

decoding identifier values for image frames according to a numbering scheme; and

resetting the identifier value for the indicated first image frame of [[an]]the independent sequence.

11. (Original) A method according to claim 10, wherein the indication is a separate flag included in the header of a slice.

12. (Canceled)

13. (Currently amended) A video decoder ~~for decoding~~comprising at least one processor and at least one memory including computer program code, the at least one memory and the

computer program code configured to, with the at least one processor, cause the video decoder to decode a compressed video sequence, wherein at least one reference image frame is predictable from at least one previous image frame, and wherein the at least one reference image frame is earlier than the at least one previous image frame in decoding order~~, the video decoder being configured to~~by causing the video decoder to at least:

decode from the video sequence an indication of at least one image frame, which is the first image frame, in decoding order, of an independent sequence, said sequence including the at least one reference image frame;

start the decoding of the video sequence from said first image frame of the independent sequence, whereby the video sequence is decoded without prediction from any image frame decoded prior to said first image frame; and

decode identifier values for image frames according to a numbering scheme; and

reset the identifier value for the indicated first image frame of [[an]]the independent sequence.

14. (Currently amended) A computer program product, stored on a non-transitory computer readable medium and executable in a data processing device, for decoding a compressed video sequence, wherein at least one reference image frame is predictable from at least one previous image frame, and wherein the at least one reference image frame is earlier than the at least one previous image frame in decoding order, the computer program product comprising:

a computer program code for decoding from the video sequence an indication of at least one image frame, which is the first image frame, in decoding order, of an independent sequence, said sequence including the at least one reference image frame;

a computer program code for starting the decoding of the video sequence from said first image frame of the independent sequence, whereby the video sequence is decoded without prediction from any image frame decoded prior to said first image frame;

a computer program code for decoding identifier values for image frames according to a numbering scheme; and

a computer program code for resetting the identifier value for the indicated first image frame of [[an]]the independent sequence.

15-16. (Canceled)

REMARKS/ARGUMENTS

The rejections presented in the Office action dated December 3, 2010 (hereinafter Office Action) have been considered. Reconsideration of the pending claims and allowance of the application in view of the present response is respectfully requested.

With respect to the objection to claims 1, 2, 4-6, 8, 10, and 11, claims 1, 5, and 9 have been amended to provide explicit antecedent basis for "the image frames" in accordance with the Examiner's suggestions. Also, claims 1, 5, 9, and 10 have been amended to provide explicit antecedent basis for the claimed independent sequence. Although not identified in the statement of objection, claims 13 and 14 have also been amended to provide explicit antecedent basis for the claimed independent sequence. These changes are believed to overcome the objection; therefore, Applicant requests that the objection be removed.

In view of the above, at least claims 1, 2, 4, 9-11, and 14 are now believed to be in condition for allowance, and Applicant accordingly requests that these claims be allowed consistent with the assertions at page seven *et seq*.

With respect to the § 101 rejections of claims 5, 6, 8, and 13, Applicant respectfully traverses. As acknowledged at page ten of the Office action, claims 5 and 13 are directed to components. While the specification also refers to embodiments involving computer readable media storing program code (*e.g.*, claims 9 and 14), the claimed encoder and decoder are devices constituting statutory subject matter. The terms "encoder" and "decoder" are well-known in the art to refer to a device or circuit. Notably, the specification illustrates structural components of the encoder (*e.g.*, compressor 108 and multiplexer 110) and decoder (*e.g.*, demultiplexer 116 and decompressor 118). Paragraph [0095] teaches examples of a decoder including a low bit rate video decoder as described in the ITU-T recommendation H.263 or H.26L. Moreover, paragraph [0096] describes positioning of the encoder and decoder that would not be possible if they were only software. However, paragraph [0097] has also been amended to remove an interpretation of the claimed encoder and decoder as being software only. Since the claimed encoder and

decoder are directed to statutory subject matter, Applicant requests that the rejection be withdrawn.

With respect to the § 112, first paragraph, rejection of claims 5, 6, 8, and 13, independent claims 5 and 13 have been amended to recite that the respective encoder and decoder comprise at least one processor and at least one memory. Support for these changes may be found in the original specification, for example, at paragraphs [0088], [0089], [0093], [0096], and [0097]; therefore, these changes do not introduce new matter. Since the device claims each recite more than a single component, the rejection is believed to be overcome. Applicant accordingly requests that the rejection be withdrawn.

In view of the above, Applicant believes each of the pending claims is in condition for allowance and requests that the application be allowed.

Authorization is given to charge Deposit Account No. 50-3581 (TPO.004.C2) any necessary fees for this filing. If the Examiner believes it necessary or helpful, the Examiner is invited to contact the undersigned attorney to discuss any issues related to this case.

Respectfully submitted,

HOLLINGSWORTH & FUNK, LLC
8500 Normandale Lake Blvd., Suite 320
Minneapolis, MN 55437
952.232.6907

Date: February 2, 2011

By: /Erin Nichols Matkaiti/

Erin Nichols Matkaiti
Reg. No. 57,125

Electronic Acknowledgement Receipt

EFS ID:	9364608
Application Number:	11338934
International Application Number:	
Confirmation Number:	6610
Title of Invention:	Grouping of image frames in video coding
First Named Inventor/Applicant Name:	Miska Hannuksela
Customer Number:	10888
Filer:	Erin M. Nichols
Filer Authorized By:	
Attorney Docket Number:	TPO.004.C2
Receipt Date:	02-FEB-2011
Filing Date:	25-JAN-2006
Time Stamp:	18:59:12
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1		OAR-TPO004C2-OA03122010.pdf	87566 573dcb95a8e974414fe4422090151ea0daabee3b	yes	9

	Multipart Description/PDF files in .zip description		
	Document Description	**Start**	**End**
	Amendment After Final	1	1
	Specification	2	2
	Claims	3	7
	Applicant Arguments/Remarks Made in an Amendment	8	9
Warnings:			
Information:			
	Total Files Size (in bytes):	87566	

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

PTO/SB/06 (07-06)
Approved for use through 1/31/2007. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

PATENT APPLICATION FEE DETERMINATION RECORD Substitute for Form PTO-875	Application or Docket Number 11/338,934	Filing Date 01/25/2006	☐ To be Mailed

APPLICATION AS FILED – PART I

	(Column 1)	(Column 2)	SMALL ENTITY ☐		OR	OTHER THAN SMALL ENTITY	
FOR	NUMBER FILED	NUMBER EXTRA	RATE ($)	FEE ($)		RATE ($)	FEE ($)
☐ BASIC FEE (37 CFR 1.16(a), (b), or (c))	N/A	N/A	N/A			N/A	
☐ SEARCH FEE (37 CFR 1.16(k), (i), or (m))	N/A	N/A	N/A			N/A	
☐ EXAMINATION FEE (37 CFR 1.16(o), (p), or (q))	N/A	N/A	N/A			N/A	
TOTAL CLAIMS (37 CFR 1.16(i))	minus 20 =	*	X $ =		OR	X $ =	
INDEPENDENT CLAIMS (37 CFR 1.16(h))	minus 3 =	*	X $ =			X $ =	
☐ APPLICATION SIZE FEE (37 CFR 1.16(s))	If the specification and drawings exceed 100 sheets of paper, the application size fee due is $250 ($125 for small entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s).						
☐ MULTIPLE DEPENDENT CLAIM PRESENT (37 CFR 1.16(j))							
* If the difference in column 1 is less than zero, enter "0" in column 2.			TOTAL			TOTAL	

APPLICATION AS AMENDED – PART II

AMENDMENT	02/02/2011	(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA	SMALL ENTITY RATE ($)	ADDITIONAL FEE ($)	OR	OTHER THAN SMALL ENTITY RATE ($)	ADDITIONAL FEE ($)
	Total (37 CFR 1.16(i))	* 11	Minus	** 20	= 0	X $ =		OR	X $52=	0
	Independent (37 CFR 1.16(h))	* 6	Minus	***6	= 0	X $ =		OR	X $220=	0
	☐ Application Size Fee (37 CFR 1.16(s))									
	☐ FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))							OR		
						TOTAL ADD'L FEE		OR	TOTAL ADD'L FEE	0

AMENDMENT		(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA	RATE ($)	ADDITIONAL FEE ($)		RATE ($)	ADDITIONAL FEE ($)
	Total (37 CFR 1.16(i))	*	Minus	**	=	X $ =		OR	X $ =	
	Independent (37 CFR 1.16(h))	*	Minus	***	=	X $ =		OR	X $ =	
	☐ Application Size Fee (37 CFR 1.16(s))									
	☐ FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))							OR		
						TOTAL ADD'L FEE		OR	TOTAL ADD'L FEE	

* If the entry in column 1 is less than the entry in column 2, write "0" in column 3.

** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 20, enter "20".

*** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 3, enter "3".

The "Highest Number Previously Paid For" (Total or Independent) is the highest number found in the appropriate box in column 1.

Legal Instrument Examiner:
/TRACIE ROBERTSON/

This collection of information is required by 37 CFR 1.16. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

OK to enter
/SL/

IN THE UNITED STATES PATENT & TRADEMARK OFFICE

Applicant:	HANNUKSELA	Examiner:	Lee, S.
Serial No.:	11/338,934	Group Art Unit:	2611
Filed:	January 25, 2006	Docket No.:	TPO.004.C2
Confirmation No.:	6610	Customer No.:	10888
Title:	GROUPING OF IMAGE FRAMES IN VIDEO CODING		

CERTIFICATE UNDER 37 CFR 1.8: The undersigned hereby certifies that this paper is being electronically transmitted by EFS-WEB to the United States Patent and Trademark Office on February 2, 2011.

By: /Erin Nichols Matkaiti/
Erin Nichols Matkaiti

AMENDMENT & RESPONSE UNDER 37 C.F.R. § 1.116

Mail Stop AF
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

In response to the Office action mailed December 3, 2010, setting a three (3) month shortened statutory period for response, Applicant submits the present amendment and response under 37 C.F.R. § 1.116.

Claims 1, 2, 4-6, 8-11, 13, and 14 remain pending in the application. Claims 1, 5, 9, 10, 13, and 14 have been amended. No claims have been added or canceled. Applicant respectfully requests reconsideration and allowance of all pending claims.

AMENDMENTS TO THE SPECIFICATION begin on page 2.

AMENDMENTS TO THE CLAIMS are reflected in the listing of claims which begins on page 3 of this paper.

REMARKS/ARGUMENTS begin on page 8.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
11/338,934	01/25/2006	Miska Hannuksela	TPO.004.C2	6610

10888 7590 03/04/2011

Hollingsworth & Funk
8500 Normandale Lake Blvd., Suite 320
Minneapolis, MN 55437

EXAMINER	
LEE, SIU M	

ART UNIT	PAPER NUMBER
2611	

NOTIFICATION DATE	DELIVERY MODE
03/04/2011	ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

tdotter@hfiplaw.com
roswood@hfiplaw.com

Office Action Summary	Application No.	Applicant(s)	
	11/338,934	HANNUKSELA, MISKA	
	Examiner	Art Unit	
	SIU M. LEE	2611	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *02 February 2011*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1,2,4-6,8-11,13 and 14* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1,2,4-6,8-11,13 and 14* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *1/25/2006* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☒ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☒ Certified copies of the priority documents have been received in Application No. 10/306942.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

DETAILED ACTION

Transitional After Final Practice

1. Since this application is eligible for the transitional procedure of 37 CFR 1.129(a), and the fee set forth in 37 CFR 1.17(r) has been timely paid, the finality of the previous Office action is hereby withdrawn pursuant to 37 CFR 1.129(a). Applicant's first submission after final filed on 12/3/2011 has been entered.

Response to Arguments

2. Applicant's arguments, see page 8-9, filed on 12/3/2011, with respect to the rejections of claims 5-8 and 13 under 35 U.S.C. § 101 and claims 5-8 and 13 under 35 U.S.C. § 112, 1st paragraph have been fully considered and are persuasive. Therefore, the rejection has been withdrawn.

Claim Rejections - 35 USC § 112

3. The following is a quotation of the second paragraph of 35 U.S.C. 112:

> The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

4. Claims 1-2, 4-6, 8-11, 13-14 are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention.

Independent claims 1, 5, 9-10, 13-14 all recite in the preamble a limitation "at least one reference image frame is predictable from at least one previous image frame,

and wherein the at least one reference image frame is earlier than the at least one previous image frame in decoding order". As the limitation states, the reference image frame is predictable from a previous image frame, therefore, it is essential for the previous image frame to be decoded before it can use to predict the reference image frame. It is unclear when the reference image frame is decoded earlier than the previous image frame, it would not be predict from the previous image frame since the reference image frame is earlier in decoding order. As stated in the abstract of the instant application, "In the decoding phase, the indication of at least one image frame is decoded from the video sequence, and the decoding of the video sequence is started from **said first image frame of the independent sequence**, whereby the video sequence is decoded without prediction from any image frame decoded prior to said first image frame"; that is the first I frame of the sequence is always being decode first. Therefore, it is unclear what "at least one reference image frame is predictable from at least one previous image frame, and wherein the at least one reference image frame is earlier than the at least one previous image frame in decoding order" means.

Independent claims 9-10, 13-14 all recite in the preamble a limitation "resetting the identifier value for the indicated first image frame of the independent sequence"; as disclose in paragraph 0018 of the publication of the instant application (2006/0120451 A1), the resetting step of the identifier value for the indicated first picture of an independent sequence is performed in the encoder, however, claims 9-11, 13-14 are

directed to a decoding, it is unclear how the decoder reset the identifier value of the indicated first image frame of the independent sequence.

Claim Rejections - 35 USC § 103

5. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

6. Claims 1 and 10 are rejected under 35 U.S.C. 103(a) as being unpatentable over Apostolopoulos (US 7,103,669 B2).

(1) Regarding claim 1:

(claim 1 is rejected without the consideration of the limitation in the preamble because of the 112 rejection as explained in above)

Apostolopoulos discloses an encoder comprising:

encoding into the video sequence an indication of at least one image frame, which is the first image frame (encoded I1 frame as shown in figure 4), in decoding order, of the independent sequence (decoding of the P frames are depending on the I1 frame), said at least one reference image frame being included in the sequence (all P frames are predict from I1 frame, the examiner interprets the claimed at least one reference image frame as one of the P frames as shown in figure 4);

encoding identifier values for the image frames according to a numbering scheme (the P frames are decoded sequentially from I1, column 9, lines 60-65; the numbering scheme of odd 1,3,5….15 and even 2,4,6…).

Apostolopoulos fails to explicitly disclose resetting the identifier value for the indicated first image frame of the independent sequence.

However, each sequence that is encoded is independent and it would have been obvious to one of ordinary skill in the art that it is not necessary to incrementing the identifier of the first image frame because it prevent using extra bit to represent a higher number of identifier and reduce the amount of data to be transmitted to the decoder.

(2) Regarding claim 10:

(claim 10 is rejected without the consideration of the limitation in the preamble because of the 112 rejection as explained in above, the examiner interprets the resetting of the identifier values are reset in the encoder)

Apostolopoulos discloses a method for decoding comprising:

decoding from the video sequence an indication of at least one image frame, which is the first image frame (the decoder decode the even frames and odd frames sequence, column 8, lines 40-49 as disclose in figure 4, the first frame is identified as an I1 frame for prediction of the following frames, column 9, lines 60-65), in decoding order, of an independent sequence, said sequence including the at least one reference image frame (all P frames are predict from I1 frame, the examiner interprets the claimed at least one reference image frame as one of the P frames as shown in figure 4);

starting the decoding of the video sequence from said first image frame of the independent sequence, whereby the video sequence is decoded without prediction from any image frame decoded prior to said first image frame (all the other P frames are depend on the I1 frame, therefore, no prediction is need from any image frame prior to I1, column 9, lines 60-65);

decoding identifier values for image frames according to a numbering scheme (the P frames are decoded sequentially from I1, column 9, lines 60-65).

Apostolopoulos fails to explicitly disclose resetting the identifier value for the indicated first image frame of the independent sequence.

However, each sequence that is encoded is independent and it would have been obvious to one of ordinary skill in the art that it is not necessary to incrementing the identifier of the first image frame because it prevent using extra bit to represent a higher number of identifier and reduce the amount of data to be transmitted to the decoder.

7. Claims 2 and 11 are rejected under 35 U.S.C. 103(a) as being unpatentable over Apostolopoulos (US 7,103,669 B2) as applied to claims 1 and 10 above, and further in view of Yagasaki et al. (US 5,786,858).

Apostolopoulos discloses all subject matter as discussed in claim 10 except wherein the indication is a separate flag included in the header of a slice.

However, Yagasaki discloses putting a flag for indicating an intra-frame coded block in the slice header (column 10, line 65 – column 11, line 2).

It is desirable to have the indication is a separate flag included in the header of a slice because it prevents a decoding process of the entire bit stream for extraction of the flag (column 11, lines 1-2). Therefore, it would have been obvious to one of ordinary skill in the art at the time of invention to employ the teaching of Yagasaki in the decoder of Apostolopoulos to simplify the extraction process of the identifier.

8. Claims 5, 9, and 13-14 are rejected under 35 U.S.C. 103(a) as being unpatentable over Apostolopoulos (US 7,103,669 B2) as applied to claim 10 above, and further in view of Langberg et al. (US 5,852,630).

(1) Regarding claims 5 and 9:

Apostolopoulos discloses an encoder comprising:

encoding into the video sequence an indication of at least one image frame, which is the first image frame (encoded I1 frame as shown in figure 4), in decoding order, of the independent sequence (decoding of the P frames are depending on the I1 frame), said at least one reference image frame being included in the sequence (all P frames are predict from I1 frame, the examiner interprets the claimed at least one reference image frame as one of the P frames as shown in figure 4);

encoding identifier values for the image frames according to a numbering scheme (the P frames are decoded sequentially from I1, column 9, lines 60-65; the numbering scheme of odd 1,3,5….15 and even 2,4,6…).

Apostolopoulos fails to explicitly disclose (a) resetting the identifier value for the indicated first image frame of the independent sequence; and (b) a computer program

product stored on a non-transitory computer readable medium and executable in a data processing device for implementing the method.

With respect to (a), each sequence that is encoded is independent and it would have been obvious to one of ordinary skill in the art that it is not necessary to incrementing the identifier of the first image frame because it prevent using extra bit to represent a higher number of identifier and reduce the amount of data to be transmitted to the decoder.

With respect to (b), Langberg et al. teaches that the method and apparatus for a transceiver warm start activation procedure with precoding can be implemented in software stored in a computer-readable medium. The computer-readable medium is an electronic, magnetic, optical, or other physical device or means that can be contain or store a computer program for use by or in connection with a computer-related system (processor) or method (column 3, lines 51-65). One skilled in the art would have clearly recognized that the method of Apostolopoulos would have been implemented in a software. The implemented software would perform same function of the hardware for less expense, adaptability, and flexibility. Therefore, it would have been obvious to have used the software as taught by Langberg et al. for implementing the decoding method of Apostolopoulos in order to reduce cost and improve the adaptability and flexibility of the encoder.

(2) Regarding claim 13 and 14:

Apostolopoulos discloses a method for decoding comprising:

decoding from the video sequence an indication of at least one image frame, which is the first image frame (the decoder decode the even frames and odd frames sequence, column 8, lines 40-49 as disclose in figure 4, the first frame is identified as an I1 frame for prediction of the following frames, column 9, lines 60-65), in decoding order, of an independent sequence, said sequence including the at least one reference image frame (all P frames are predict from I1 frame, the examiner interprets the claimed at least one reference image frame as one of the P frames as shown in figure 4);

starting the decoding of the video sequence from said first image frame of the independent sequence, whereby the video sequence is decoded without prediction from any image frame decoded prior to said first image frame (all the other P frames are depend on the I1 frame, therefore, no prediction is need from any image frame prior to I1, column 9, lines 60-65);

decoding identifier values for image frames according to a numbering scheme (the P frames are decoded sequentially from I1, column 9, lines 60-65).

Apostolopoulos fails to explicitly disclose (a) resetting the identifier value for the indicated first image frame of the independent sequence; and (b) the decoder comprising at least one processor and at leas tone memory including computer program code to perform the method.

With respect to (a), each sequence that is encoded is independent and it would have been obvious to one of ordinary skill in the art that it is not necessary to incrementing the identifier of the first image frame because it prevent using extra bit to

represent a higher number of identifier and reduce the amount of data to be transmitted to the decoder.

with respect to (b), Langberg et al. teaches that the method and apparatus for a transceiver warm start activation procedure with precoding can be implemented in software stored in a computer-readable medium. The computer-readable medium is an electronic, magnetic, optical, or other physical device or means that can be contain or store a computer program for use by or in connection with a computer-related system (processor) or method (column 3, lines 51-65). One skilled in the art would have clearly recognized that the method of Apostolopoulos would have been implemented in a software. The implemented software would perform same function of the hardware for less expense, adaptability, and flexibility. Therefore, it would have been obvious to have used the software as taught by Langberg et al. for implementing the decoding method of Apostolopoulos in order to reduce cost and improve the adaptability and flexibility of the decoder.

9. Claim 2 is rejected under 35 U.S.C. 103(a) as being unpatentable over Apostolopoulos (US 7,103,669 B2) as applied to claim 10 above, and further in view of Yagasaki et al. (US 5,786,858).

Apostolopoulos discloses all subject matter as discussed in claim 10 except wherein the indication is a separate flag included in the header of a slice.

However, Yagasaki discloses putting a flag for indicating an intra-frame coded block in the slice header (column 10, line 65 – column 11, line 2).

It is desirable to have the indication is a separate flag included in the header of a slice because it prevents a decoding process of the entire bit stream for extraction of the flag (column 11, lines 1-2). Therefore, it would have been obvious to one of ordinary skill in the art at the time of invention to employ the teaching of Yagasaki in the decoder of Apostolopoulos to simplify the extraction process of the identifier.

10. Claim 6 is rejected under 35 U.S.C. 103(a) as being unpatentable over Apostolopoulos (US 7,103,669 B2) in view of Langberg et al. (US 5,852,630) as applied to claim 5 above, and further in view of Yagasaki et al. (US 5,786,858).

Apostolopoulos and Langberg discloses all subject matter as discussed in claim 5 except wherein the indication is a separate flag included in the header of a slice.

However, Yagasaki discloses putting a flag for indicating an intra-frame coded block in the slice header (column 10, line 65 – column 11, line 2).

It is desirable to have the indication is a separate flag included in the header of a slice because it prevents a decoding process of the entire bit stream for extraction of the flag (column 11, lines 1-2). Therefore, it would have been obvious to one of ordinary skill in the art at the time of invention to employ the teaching of Yagasaki in the decoder of Apostolopoulos and Langberg to simplify the extraction process of the identifier.

Conclusion

11. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

Hasebe et al. (US 6,483,875 B1) discloses a picture signal processing apparatus.

Shinoda (US 5,680,322) discloses a method and apparatus for dynamic image data transmission.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to SIU M. LEE whose telephone number is (571)270-1083. The examiner can normally be reached on Mon-Fri, 7:30-4:00 with every other Friday off.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Chieh Fan can be reached on (571) 272-3042. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Siu M Lee/
Examiner, Art Unit 2611
2/28/2011

Notice of References Cited	Application/Control No. 11/338,934	Applicant(s)/Patent Under Reexamination HANNUKSELA, MISKA	
	Examiner SIU M. LEE	Art Unit 2611	Page 1 of 1

U.S. PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Name	Classification
*	A	US-5,680,322	10-1997	Shinoda, Mayumi	714/18
*	B	US-6,483,875	11-2002	Hasebe et al.	375/240.15
*	C	US-5,786,858	07-1998	Yagasaki et al.	375/240.15
*	D	US-7,103,669	09-2006	Apostolopoulos, John G.	709/231
*	E	US-5,852,630	12-1998	Langberg et al.	375/219
	F	US-			
	G	US-			
	H	US-			
	I	US-			
	J	US-			
	K	US-			
	L	US-			
	M	US-			

FOREIGN PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Country	Name	Classification
	N					
	O					
	P					
	Q					
	R					
	S					
	T					

NON-PATENT DOCUMENTS

*		Include as applicable: Author, Title Date, Publisher, Edition or Volume, Pertinent Pages)
	U	
	V	
	W	
	X	

*A copy of this reference is not being furnished with this Office action. (See MPEP § 707.05(a).)
Dates in MM-YYYY format are publication dates. Classifications may be US or foreign.

Index of Claims	Application/Control No.	Applicant(s)/Patent Under Reexamination
	11338934	HANNUKSELA, MISKA
	Examiner	Art Unit
	SIU M LEE	2611

✓	Rejected	-	Cancelled	N	Non-Elected	A	Appeal
=	Allowed	÷	Restricted	I	Interference	O	Objected

☐ Claims renumbered in the same order as presented by applicant ☐ CPA ☐ T.D. ☐ R.1.47

CLAIM		DATE								
Final	Original	05/21/2010	12/02/2010	02/28/2011						
	1	✓	O	✓						
	2	✓	O	✓						
	3	✓	-	-						
	4	✓	O	✓						
	5	✓	✓	✓						
	6	✓	✓	✓						
	7	✓	-	-						
	8	✓	✓	✓						
	9	✓	O	✓						
	10	✓	O	✓						
	11	✓	O	✓						
	12	✓	-	-						
	13	✓	✓	✓						
	14	✓	O	✓						
	15	✓	-	-						
	16	✓	-	-						

Search Notes	Application/Control No.	Applicant(s)/Patent Under Reexamination
	11338934	HANNUKSELA, MISKA
	Examiner	Art Unit
	SIU M LEE	2611

SEARCHED

Class	Subclass	Date	Examiner
375	240.12, 240.25	5/21/2010	Siu M. Lee

SEARCH NOTES

Search Notes	Date	Examiner
EAST	5/21/2010	Siu M. Lee
Inventor name search - Miska hannuksela	5/21/2010	Siu M. Lee
Discuss the application with Chieh Fan	5/21/2010	Siu M. Lee
EAST	12/1/2010	Siu M. Lee
Discuss the amendment with Chieh Fan	12/1/2010	Siu M. Lee
Discuss the 101 and 112 rejection of claims 5-8 and 13 with John Peng	12/1/2010	Siu M. Lee
EAST	2/28/2011	Siu M. Lee
Discuss the amendment with regard to 101 problem with Daniel Swerdlow	2/25/2011	Siu M. Lee

INTERFERENCE SEARCH

Class	Subclass	Date	Examiner

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

BIB DATA SHEET

CONFIRMATION NO. 6610

SERIAL NUMBER	FILING or 371(c) DATE RULE	CLASS	GROUP ART UNIT	ATTORNEY DOCKET NO.
11/338,934	01/25/2006	375	2611	TPO.004.C2

APPLICANTS
Miska Hannuksela, Tampere, FINLAND;

**** CONTINUING DATA ***************************
This application is a CON of 10/306,942 11/29/2002 PAT 7,894,521

**** FOREIGN APPLICATIONS ***************************
FINLAND 20020127 01/23/2002

**** IF REQUIRED, FOREIGN FILING LICENSE GRANTED ****
02/28/2006

Foreign Priority claimed ☑ Yes ☐ No 35 USC 119(a-d) conditions met ☑ Yes ☐ No Verified and Acknowledged /SIU M LEE/ Examiner's Signature	☐ Met after Allowance Initials	STATE OR COUNTRY	SHEETS DRAWINGS	TOTAL CLAIMS	INDEPENDENT CLAIMS
		FINLAND	4	16	7

ADDRESS

Hollingsworth & Funk
8500 Normandale Lake Blvd., Suite 320
Minneapolis, MN 55437
UNITED STATES

TITLE

Grouping of image frames in video coding

FILING FEE RECEIVED		
1800	FEES: Authority has been given in Paper No._____ to charge/credit DEPOSIT ACCOUNT No._____ for following:	☐ All Fees ☐ 1.16 Fees (Filing) ☐ 1.17 Fees (Processing Ext. of time) ☐ 1.18 Fees (Issue) ☐ Other _____ ☐ Credit

BIB (Rev. 05/07).

EAST Search History

EAST Search History (Prior Art)

Ref #	Hits	Search Query	DBs	Default Operator	Plurals	Time Stamp
L1	2	encod$3 with image frame with decoding near2 order	US-PGPUB; USPAT; USOCR	ADJ	ON	2011/02/28 07:10
L2	358	encod$3 with image frame with decoding	US-PGPUB; USPAT; USOCR	ADJ	ON	2011/02/28 07:10
L3	22	encod$3 with image frame with decoding with order	US-PGPUB; USPAT; USOCR	ADJ	ON	2011/02/28 07:10
L4	18	3 and (@pd<="20020123" OR @ad<="20020123" OR @rlad<="20020123" OR @prad<="20020123")	US-PGPUB; USPAT; USOCR	ADJ	ON	2011/02/28 07:10
L5	4	image frame and (header with indicat$4 with decod$3 order)	US-PGPUB; USPAT; USOCR	ADJ	ON	2011/02/28 07:13
L6	2	image frame and (header with bit with decod$3 order)	US-PGPUB; USPAT; USOCR	ADJ	ON	2011/02/28 07:16
L7	4	image frame and (header with (indicat$3 or identif$4) with decod$3 order)	US-PGPUB; USPAT; USOCR	ADJ	ON	2011/02/28 07:17
L8	1	"5786858".pn. and ff	US-PGPUB; USPAT; USOCR	ADJ	ON	2011/02/28 07:18
L9	1	"5786858".pn. and ff???	US-PGPUB; USPAT; USOCR	ADJ	ON	2011/02/28 07:19
L10	314	random access with image frame	US-PGPUB; USPAT; USOCR	ADJ	ON	2011/02/28 07:26
L11	29	10 and header	US-PGPUB; USPAT; USOCR	ADJ	ON	2011/02/28 07:27
L12	23	11 and (@pd<="20020123" OR @ad<="20020123" OR @rlad<="20020123" OR @prad<="20020123")	US-PGPUB; USPAT; USOCR	ADJ	ON	2011/02/28 07:27
L13	18	header with (code or bit or indicat$3 or identif$5) with decod$3 order	US-PGPUB; USPAT; USOCR	ADJ	ON	2011/02/28 07:37
L14	7	13 and (@pd<="20020123" OR @ad<="20020123" OR @rlad<="20020123" OR @prad<="20020123")	US-PGPUB; USPAT; USOCR	ADJ	ON	2011/02/28 07:38
S2	1	decod$3 with flag with header with intra frame	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2011/02/25 14:11

S3	25	("4882732" | "4930024" | "5057949" | "5126852" | "5136391" | "5140437" | "5231484" | "5247363" | "5260783" | "5282049" | "5305113" | "5371602" | "5377051" | "5455684" | "5543846").PN. OR ("5786858").URPN.	US-PGPUB; USPAT; USOCR	ADJ	ON	2011/02/25 14:12
S4	22	S3 and (@pd<="20020123" OR @ad<="20020123" OR @rlad<="20020123" OR @prad<="20020123")	US-PGPUB; USPAT; USOCR	ADJ	ON	2011/02/25 14:14
S5	3	"7103669".pn.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2011/02/26 18:36
S6	15	(flag or indicator) with header with image frame	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2011/02/26 22:25
S7	8	S6 and (@pd<="20020123" OR @ad<="20020123" OR @rlad<="20020123" OR @prad<="20020123")	US-PGPUB; USPAT; USOCR	ADJ	ON	2011/02/26 22:25
S8	1658	flag with slice	US-PGPUB; USPAT; USOCR	ADJ	ON	2011/02/26 23:00
S9	383	S8 and (@pd<="20020123" OR @ad<="20020123" OR @rlad<="20020123" OR @prad<="20020123")	US-PGPUB; USPAT; USOCR	ADJ	ON	2011/02/26 23:00
S10	73	S9 and intra frame	US-PGPUB; USPAT; USOCR	ADJ	ON	2011/02/26 23:01
S11	5	S9 and (intra?frame with flag with slice)	US-PGPUB; USPAT; USOCR	ADJ	ON	2011/02/26 23:01
S12	25	("4882732" | "4930024" | "5057949" | "5126852" | "5136391" | "5140437" | "5231484" | "5247363" | "5260783" | "5282049" | "5305113" | "5371602" | "5377051" | "5455684" | "5543846").PN. OR ("5786858").URPN.	US-PGPUB; USPAT; USOCR	ADJ	ON	2011/02/26 23:02
S13	22	S12 and (@pd<="20020123" OR @ad<="20020123" OR @rlad<="20020123" OR @prad<="20020123")	US-PGPUB; USPAT; USOCR	ADJ	ON	2011/02/26 23:02

S14	1	11/338934 and ((ref or reference) with previous) and (decod$3 near3 order)	US-PGPUB; USPAT; USOCR	ADJ	ON	2011/02/27 07:01
S15	1	"7103669".pn.	US-PGPUB; USPAT; USOCR	ADJ	ON	2011/02/27 09:03
S16	2	11/338934 and (previous)	US-PGPUB; USPAT; USOCR	ADJ	ON	2011/02/27 17:54
S17	2	11/338934 and (decoding order)	US-PGPUB; USPAT; USOCR	ADJ	ON	2011/02/27 18:14
S18	2	11/338934 and (decod$3 order)	US-PGPUB; USPAT; USOCR	ADJ	ON	2011/02/27 18:19
S19	2	11/338934 and (decod$3)	US-PGPUB; USPAT; USOCR	ADJ	ON	2011/02/27 18:20
S20	1	11/338934 and (reset$4)	US-PGPUB; USPAT; USOCR	ADJ	ON	2011/02/27 18:29
S21	5	reset$4 with I frame with (identif $4 or inducat$3)	US-PGPUB; USPAT; USOCR	ADJ	ON	2011/02/27 18:35

EAST Search History (Interference)

<This search history is empty>

2/28/2011 8:23:08 AM
C:\Documents and Settings\slee7\My Documents\EAST\Workspaces\11338934 201102.wsp

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	PATENT NUMBER	GROUP ART UNIT	FILE WRAPPER LOCATION
11/338,934		2611	

OC000000046476016

Correspondence Address/Fee Address Change

The following fields have been set to Customer Number 10888 on 11/30/2010

- **Correspondence Address**
- **Power of Attorney Address**

The address of record for Customer Number 10888 is:

10888
Hollingsworth & Funk
8500 Normandale Lake Blvd., Suite 320
Minneapolis, MN 55437

FORM 1449* SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT IN AN APPLICATION (Use several sheets if necessary)	Docket Number: TPO.004.C2	Application Number: 11/338,934
	Applicant: HANNUKSELA	
	Filing Date: 01/25/2006	Group Art Unit: 2611

U.S. PATENT DOCUMENTS

EXAMINER INITIAL	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE

FOREIGN PATENT DOCUMENTS

	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	TRANSLATION	
						YES	NO

OTHER DOCUMENTS (Including Author, Title, Date, Pertinent Pages, Etc.)

	2002	Hannuksela, "Signaling of Enhanced GOPs", Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, 4th Meeting, Austria, July 22-26, 2002.
	4/5/2011	File History for EP Application No. 03700321.7 as retrieved from the European Patent Office Electronic File System on 4/5/2011, 222 pages.

EXAMINER	DATE CONSIDERED
EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609; draw line through citation if not in conformance and not considered. Include copy of this form for next communication to the Applicant.	

*Substitute Disclosure Statement Form (PTO-1449) Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

IN THE UNITED STATES PATENT & TRADEMARK OFFICE

Applicant:	HANNUKSELA	Examiner:	Lee, S.
Serial No.:	11/338,934	Group Art Unit:	2611
Filed:	January 25, 2006	Docket No.:	TPO.004.C2
Confirmation No.:	6610	Customer No.:	10888
Title:	GROUPING OF IMAGE FRAMES IN VIDEO CODING		

SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT
UNDER 37 C.F.R. § 1.97(c)(1)

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

With regard to the above-identified application, the items of information listed on the enclosed Form 1449 are brought to the attention of the Examiner.

This statement should be considered because it is submitted after the mailing date of the first Office action on the merits but before the mailing date of a final Office action under 37 C.F.R. § 1.113 or a Notice of Allowance under 37 C.F.R. § 1.311.

The items listed on the enclosed Form 1449 were first cited in a communication from a foreign patent office regarding a counterpart foreign application not more than three months prior to the filing of the information disclosure statement. Accordingly, no fee is due for consideration of the items listed on the enclosed Form 1449. However, if fees are deemed necessary, authorization is given to charge Deposit Account 50-3581 (TPO.004.C2) for consideration of the information disclosure statement.

In accordance with 37 C.F.R. § 1.98(a)(2), a copy of the non-U.S. patent/application documents listed on the enclosed Form 1449 is provided.

Please note that any notations or markings on the attached documents do not reflect particular relevance, or lack thereof, to the present application, nor were they necessarily made by anyone affiliated with the prosecution of the present application.

No representation is made that the reference is "prior art" within the meaning of 35 U.S.C. §§ 102 and 103. In addition, Applicant does not represent that a reference has been thoroughly reviewed or that any relevance of any portion of a reference is intended, and reserves the right to establish otherwise under 37 C.F.R. § 1.131 or others.

Consideration of the items listed is respectfully requested. According to MPEP § 609, Applicant requests that the Examiner return a copy of the attached Form 1449, marked as being considered and initialed by the Examiner, to the undersigned with the next official communication.

Respectfully submitted,
HOLLINGSWORTH & FUNK, LLC
8500 Normandale Lake Blvd., Suite 320
Minneapolis, MN 55437
952.854.2700

By: /Erin Nichols Matkaiti/
Erin Nichols Matkaiti
Reg. No. 57,125

Electronic Acknowledgement Receipt

EFS ID:	9810103
Application Number:	11338934
International Application Number:	
Confirmation Number:	6610
Title of Invention:	Grouping of image frames in video coding
First Named Inventor/Applicant Name:	Miska Hannuksela
Customer Number:	10888
Filer:	Erin M. Nichols
Filer Authorized By:	
Attorney Docket Number:	TPO.004.C2
Receipt Date:	05-APR-2011
Filing Date:	25-JAN-2006
Time Stamp:	19:03:48
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Information Disclosure Statement (IDS) Filed (SB/08)	1449-13-TPO004C2.pdf	59869 5b8deecba38c717dbf226b58d991b55531b5bf7f	no	1

Warnings:	
Information:	

This is not an USPTO supplied IDS fillable form

2	NPL Documents	FH-EP037003217.pdf	11489599 eb3fbc40a44f1218f89d4b32e1668595b2498987	no	222
Warnings:					
Information:					
3	NPL Documents	Hannuksela-2.pdf	73745 d1af007096e859d2cd6d5cc1d112dbf8df536262	no	9
Warnings:					
Information:					
4	Transmittal Letter	IDS-13-TPO004C2.pdf	56596 dc8f81fd9b533eb00b878aef5680652538bf003b	no	2
Warnings:					
Information:					
		Total Files Size (in bytes):	11679809		

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111

If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371

If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office

If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

IN THE UNITED STATES PATENT & TRADEMARK OFFICE

Applicant:	HANNUKSELA	Examiner:	Lee, S.
Serial No.:	11/338,934	Group Art Unit:	2611
Filed:	January 25, 2006	Docket No.:	TPO.004.C2
Confirmation No.:	6610	Customer No.:	10888
Title:	GROUPING OF IMAGE FRAMES IN VIDEO CODING		

CERTIFICATE UNDER 37 CFR 1.8: The undersigned hereby certifies that this paper is being electronically transmitted by EFS-WEB to the United States Patent and Trademark Office on June 2, 2011.

By: /Erin Nichols Matkaiti/
Erin Nichols Matkaiti

AMENDMENT & RESPONSE UNDER 37 C.F.R. § 1.111

Mail Stop Amendment
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

In response to the Office action mailed March 4, 2011, setting a three (3) month shortened statutory period for response, Applicant submits the present amendment and response under 37 C.F.R. § 1.111.

Claims 1, 2, 4-6, 8-11, 13, and 14 remain pending in the application. Claims 1, 5, 9, 10, 13, and 14 have been amended. No claims have been added or canceled. Applicant respectfully requests reconsideration and allowance of all pending claims.

AMENDMENTS TO THE CLAIMS are reflected in the listing of claims which begins on page 2 of this paper.

REMARKS/ARGUMENTS begin on page 7.

AMENDMENTS TO THE CLAIMS

This listing of claims replaces all prior versions of claims in the application.

1. (Currently amended) A method for encoding a video sequence comprising an independent sequence of image frames, wherein ~~at least one reference image frame is predictable from at least one previous image frame, and wherein the at least one reference image frame is earlier than the at least one previous image frame in decoding order~~all motion-compensated temporal prediction references of the independent sequence refer only to image frames within said independent sequence, the method comprising:

encoding into the video sequence an indication of at least one image frame, which is the first image frame, in decoding order, of the independent sequence~~, said at least one reference image frame being included in the sequence~~;

encoding identifier values for the image frames according to a numbering scheme; and

resetting the identifier value for the indicated first image frame of the independent sequence.

2. (Original) A method according to claim 1, further comprising:

encoding the indication into the video sequence as a separate flag included in the header of a slice.

3. (Canceled)

4. (Original) A method according to claim 1, further comprising:

encoding into the video sequence an identifier value for the independent sequence.

5. (Currently amended) A video encoder comprising at least one processor and at least one memory including computer program code, the at least one memory and the computer program code configured to, with the at least one processor, cause the video encoder to

encode a video sequence comprising an independent sequence of image frames, wherein ~~at least one reference image frame is predictable from at least one previous image frame, and wherein the at least one reference frame is earlier than the at least one previous image frame in decoding order~~all motion-compensated temporal prediction references of the independent sequence refer only to image frames within said independent sequence by causing the video encoder to at least:

encode into the video sequence an indication of at least one image frame, which is the first image frame, in decoding order, of the independent sequence~~, said at least one reference image frame being included in the sequence~~;

encode identifier values for the image frames according to a numbering scheme; and

reset the identifier value for the indicated first image frame of the independent sequence.

6. (Previously presented) A video encoder according to claim 5, wherein the video encoder is configured to

encode the indication into the video sequence as a separate flag included in the header of a slice.

7. (Canceled)

8. (Previously presented) A video encoder according to claim 5, wherein the video encoder is configured to

encode into the video sequence an identifier value for the independent sequence.

9. (Currently amended) A computer program product, stored on a non-transitory computer readable medium and executable in a data processing device, for encoding a video sequence comprising an independent sequence of image frames, wherein ~~at least one reference image frame is predictable from at least one previous image frame and wherein the at least one reference frame is earlier than the at least one previous image frame in decoding order~~all

motion-compensated temporal prediction references of the independent sequence refer only to image frames within said independent sequence, the computer program product comprising:

a computer program code for encoding into the video sequence an indication of at least one image frame, which is the first image frame, in decoding order, of the independent sequence~~, said at least one reference image frame being included in the sequence~~;

a computer code for encoding identifier values for the image frames according to a numbering scheme; and

a computer code for resetting the identifier value for the indicated first image frame of the independent sequence.

10. (Currently amended) A method for decoding a compressed video sequence, ~~wherein at least one reference image frame is predictable from at least one previous image frame and wherein the at least one reference frame is earlier than the at least one previous image frame in decoding order,~~ the method comprising:

decoding from the video sequence an indication of at least one image frame, which is the first image frame, in decoding order, of an independent sequence, [[said]]wherein all motion-compensated temporal prediction references of the independent sequence ~~including the at least one reference~~refer only to image ~~frame~~frames within said independent sequence;

starting the decoding of the video sequence from said first image frame of the independent sequence, whereby the video sequence is decoded without prediction from any image frame decoded prior to said first image frame;

decoding identifier values for image frames according to a numbering scheme; and

resetting the identifier value for the indicated first image frame of the independent sequence.

11. (Original) A method according to claim 10, wherein the indication is a separate flag included in the header of a slice.

12. (Canceled)

13. (Currently amended) A video decoder comprising at least one processor and at least one memory including computer program code, the at least one memory and the computer program code configured to, with the at least one processor, cause the video decoder to decode a compressed video sequence, ~~wherein at least one reference image frame is predictable from at least one previous image frame, and wherein the at least one reference image frame is earlier than the at least one previous image frame in decoding order~~ by causing the video decoder to at least:

decode from the video sequence an indication of at least one image frame, which is the first image frame, in decoding order, of an independent sequence, [[said]]wherein all motion-compensated temporal prediction references of the independent sequence ~~including the at least one reference~~refer only to image ~~frame~~frames within said independent sequence;

start the decoding of the video sequence from said first image frame of the independent sequence, whereby the video sequence is decoded without prediction from any image frame decoded prior to said first image frame; and

decode identifier values for image frames according to a numbering scheme; and

reset the identifier value for the indicated first image frame of the independent sequence.

14. (Currently amended) A computer program product, stored on a non-transitory computer readable medium and executable in a data processing device, for decoding a compressed video sequence, ~~wherein at least one reference image frame is predictable from at least one previous image frame, and wherein the at least one reference image frame is earlier than the at least one previous image frame in decoding order,~~ the computer program product comprising:

a computer program code for decoding from the video sequence an indication of at least one image frame, which is the first image frame, in decoding order, of an independent sequence, [[said]]wherein all motion-compensated temporal prediction references of the

independent sequence ~~including the at least one reference~~refer only to image ~~frame~~frames within said independent sequence;

a computer program code for starting the decoding of the video sequence from said first image frame of the independent sequence, whereby the video sequence is decoded without prediction from any image frame decoded prior to said first image frame;

a computer program code for decoding identifier values for image frames according to a numbering scheme; and

a computer program code for resetting the identifier value for the indicated first image frame of the independent sequence.

15-16. (Canceled)

REMARKS/ARGUMENTS

The reference to "Transitional After Final Practice" on page two of the Office action appears to be in error. Applicant notes that this application does not meet the timing requirements for Rule 1.129(a) and no such fee was paid. Applicant requests that these comments be withheld from further actions or that an explanation be provided.

The rejections presented in the Office action dated March 4, 2011 (hereinafter "Office action") have been considered. Reconsideration of the pending claims and allowance of the application in view of the present response is respectfully requested.

With respect to the § 112, second paragraph, rejection of the independent claims, each of the independent claims has been amended to remove the objected-to language. The language has been rewritten to characterize that "all motion-compensated temporal prediction references of the independent sequence refer only to image frames within said independent sequence". Since these changes merely re-characterize the original claim language, the changes do not introduce new matter. With particular respect to claims 9, 10, 13, and 14, Applicant notes that the identified language directed to "resetting" is located in the bodies, not the preambles, of these claims. In contrast to the assertion, claim 9 is directed to encoding. Moreover, discussion of these elements in the decoder may be found in the original specification, for example, at paragraph [0057] (using the numbering of U.S. Publication No. 2006/0120451). In view of the above, the claims particularly point out and distinctly claim the intended subject matter, and Applicant requests that the rejection be withdrawn.

Applicant respectfully traverses each of the §103(a) rejections based primarily upon the teachings of U.S. Patent No. 7,103,669 to Apostolopoulos (hereinafter "Apostolopoulos") because Apostolopoulos does not teach or suggest each of the claimed elements. For example, Apostolopoulos does not teach or suggest at least encoding into the video sequence, or decoding, an indication of at least one image frame, which is the first image frame, in decoding order, of the independent sequence, as claimed in each of the independent claims. Notably, Apostolopoulos does not discuss browsing a video file from a

random point and the discontinuities typically occurring in the image numbering due to such random access.

The present invention is directed to signaling to a decoder which video frame is to be decoded first as part of an independently decodable sequence of image frames, especially if there is another I-frame in the middle of the independently decodable sequence. This is discussed in paragraphs [0012-0013, 0056] (using the numbering of U.S. Publication No. 2006/0120451). Paragraph [0057] describes that there is encoded an indication of at least one image frame, which is the first image frame, in decoding order, of the independent sequence (GOP), wherein all motion-compensated temporal prediction references of the independent sequence refer only to image frames within said independent sequence, identifier values for the image frames are encoded according to a numbering scheme, and the identifier value for the indicated first image frame of the independent sequence is reset at zero.

In contrast, Apostolopoulos does not disclose or address such issues. In Apostolopoulos, a video stream is divided in two video streams, even frames encoded in a first stream and odd frame encoded in a second stream. The two streams are transmitted via separate transmission paths, decoded at reception and then combined into a complete video stream. Apostolopoulos addresses when a video file that is to be reproduced is detected to contain an error that interrupts the playback or requires the playback to be resumed from a point following the error. However, to solve this problem, Apostolopoulos suggests that if a frame is missing at the decoder, the decoding of that particular video stream is halted until the next I-frame of the video stream is received (Col. 9, line 60 – Col. 10, line 8). Thus, there is no indication, as claimed. Rather, if such an indication were included in the video stream, the decoder would immediately know where to start the decoding.

In addition, the Examiner acknowledges that Apostolopoulos does not disclose "resetting the identifier value for the indicated first image frame of the independent sequence". The assertion that such resetting would be obvious for a skilled person when aiming to prevent using an extra bit to represent higher frame numbers, is entirely unsupported. Moreover, the purpose of resetting the identifier value is not to reduce the

amount of data to be transmitted to the decoder, but to (in combination with the indication) allow the decoder to interpret the image numbers correctly and to find the correct image frame that initiates the sub-sequence from the video sequence image frames (previously cited paragraph [0057]). Without a presentation of correspondence to each of the claimed elements, each of the § 103(a) rejections is improper.

By similar reasoning, the dependent claims are also not obvious in view of the asserted references, whether each reference is considered alone or the references are considered in combination. Applicant respectfully requests withdrawal of the rejections of claims 1, 2, 4-6, 8-11, 13, and 14 and submits that all claims are in condition for allowance.

To the extent that the current response does not respond to any characterization in the Office action of the asserted art or of the claimed subject matter, or to any application in the Office action of the asserted art to any claimed subject matter, it is stated for the record that any such lack of response should not be interpreted as an acquiescence to such characterizations or applications. A detailed discussion of each of the Examiner's characterizations, or any other assertions or statements beyond that provided above is unnecessary in view of the present response. The right to address in detail any such assertions or statements in the future is reserved.

Authorization is given to charge Deposit Account No. 50-3581 (TPO.004.C2) any necessary fees for this filing. If the Examiner believes it necessary or helpful, the Examiner is invited to contact the undersigned attorney to discuss any issues related to this case.

Respectfully submitted,

HOLLINGSWORTH & FUNK, LLC
8500 Normandale Lake Blvd., Suite 320
Minneapolis, MN 55437
952.232.6907

Date: June 2, 2011

By: /Erin Nichols Matkaiti/

Erin Nichols Matkaiti
Reg. No. 57,125

IN THE UNITED STATES PATENT & TRADEMARK OFFICE

Applicant:	HANNUKSELA	Examiner:	Lee, S.
Serial No.:	11/338,934	Group Art Unit:	2611
Filed:	January 25, 2006	Docket No.:	TPO.004.C2
Confirmation No.:	6610	Customer No.:	10888
Title:	GROUPING OF IMAGE FRAMES IN VIDEO CODING		

SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT
UNDER 37 C.F.R. § 1.97(c)

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

Applicant respectfully submits the items of information on the enclosed Form 1449 for the attention of the Examiner in the above-identified application.

This statement should be considered because it is submitted after the mailing date of the first Office action on the merits but before the mailing date of a final Office action under 37 C.F.R. § 1.113 or a Notice of Allowance under 37 C.F.R. § 1.311. Please charge Deposit Account 50-3581 (TPO.004.C2) the amount of $180.00 under 37 C.F.R. § 1.17(p) for consideration of the items.

In accordance with 37 C.F.R. § 1.98(a)(2), only a copy of each foreign document or non-U.S. patent/application listed on the enclosed Form 1449 is provided.

Please note that any notations or markings on the attached documents do not reflect particular relevance, or lack thereof, to the present application, nor were they necessarily made by anyone affiliated with the prosecution of the present application.

No representation is made that a reference is "prior art" within the meaning of 35 U.S.C. §§ 102 and 103. In addition, Applicant does not represent that a reference has been thoroughly reviewed or that any relevance of any portion of a reference is intended, and reserves the right to establish otherwise under 37 C.F.R. § 1.131 or others.

Consideration of the items listed is respectfully requested. According to MPEP § 609, Applicant requests that the Examiner return a copy of the attached Form 1449, marked as being considered and initialed by the Examiner, to the undersigned with the next official communication.

Respectfully submitted,

HOLLINGSWORTH & FUNK, LLC
8500 Normandale Lake Blvd., Suite 320
Minneapolis, MN 55437
952.232.6907

By: /Erin Nichols Matkaiti/
Erin Nichols Matkaiti
Reg. No. 57,125

FORM 1449* INFORMATION DISCLOSURE STATEMENT IN AN APPLICATION (Use several sheets if necessary)	Docket Number: TPO.004.C2	Application Number: 11/338,934
	Applicant: HANNUKSELA	
	Filing Date: 1/25/2006	Group Art Unit: 2611

U.S. PATENT DOCUMENTS

EXAMINER INITIAL	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE

FOREIGN PATENT DOCUMENTS

	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	TRANSLATION	
						YES	NO

OTHER DOCUMENTS (Including Author, Title, Date, Pertinent Pages, Etc.)

	2/18/2011	Office Action dated 2/18/2011 from U.S. Serial No. 11/338,996, 18 pages.
	5/17/2011	Office Action Response dated 5/17/2011 from U.S. Serial No. 11/338,996, 9 pages.

EXAMINER	DATE CONSIDERED
EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609; draw line through citation if not in conformance and not considered. Include copy of this form for next communication to the Applicant.	

*Substitute Disclosure Statement Form (PTO-1449) Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Electronic Patent Application Fee Transmittal

Application Number:	11338934
Filing Date:	25-Jan-2006
Title of Invention:	Grouping of image frames in video coding
First Named Inventor/Applicant Name:	Miska Hannuksela
Filer:	Erin M. Nichols
Attorney Docket Number:	TPO.004.C2

Filed as Large Entity

Utility under 35 USC 111(a) Filing Fees

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Basic Filing:				
Pages:				
Claims:				
Miscellaneous-Filing:				
Petition:				
Patent-Appeals-and-Interference:				
Post-Allowance-and-Post-Issuance:				
Extension-of-Time:				

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Miscellaneous:				
Submission- Information Disclosure Stmt	1806	1	180	180
		Total in USD ($)		**180**

Electronic Acknowledgement Receipt

EFS ID:	10216166
Application Number:	11338934
International Application Number:	
Confirmation Number:	6610
Title of Invention:	Grouping of image frames in video coding
First Named Inventor/Applicant Name:	Miska Hannuksela
Customer Number:	10888
Filer:	Erin M. Nichols
Filer Authorized By:	
Attorney Docket Number:	TPO.004.C2
Receipt Date:	02-JUN-2011
Filing Date:	25-JAN-2006
Time Stamp:	13:21:50
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	yes
Payment Type	Deposit Account
Payment was successfully received in RAM	$180
RAM confirmation Number	13215
Deposit Account	503581
Authorized User	

The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows:

Charge any Additional Fees required under 37 C.F.R. Section 1.21 (Miscellaneous fees and charges)

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1		OAR-TPO004C2-OA04032011.pdf	83410 099487a97bfd6262ec23078255dae2f97c503adb	yes	9

Multipart Description/PDF files in .zip description

Document Description	Start	End
Amendment/Req. Reconsideration-After Non-Final Reject	1	1
Claims	2	6
Applicant Arguments/Remarks Made in an Amendment	7	9

Warnings:

Information:

2	Transmittal Letter	IDS-14-TPO004C2.pdf	56107 4edd1bd0b511ff1d6b1ffab066d6d14388288b44	no	2

Warnings:

Information:

3	Information Disclosure Statement (IDS) Form (SB08)	1449-14-TPO004C2.pdf	52216 19716cc11fd61404e57b52456912382b253711de	no	1

Warnings:

Information:

This is not an USPTO supplied IDS fillable form

4	Non Patent Literature	OA-180202011.pdf	610274 23af7b687ed78aed21d21653afa8dda4fd02ba4d	no	18

Warnings:

Information:

5	Non Patent Literature	OAR-TPO004C1-OA18022011.pdf	79696 47300f709550cb902f658bd15bd927096f7e4390	no	9

Warnings:

Information:

6	Fee Worksheet (SB06)	fee-info.pdf	30217 ecad26660c9a3db990345b382663d3099b9bdf94	no	2

Warnings:

Information:

Total Files Size (in bytes):	911920

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111

If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371

If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office

If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

PTO/SB/06 (07-06)
Approved for use through 1/31/2007. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

PATENT APPLICATION FEE DETERMINATION RECORD

Substitute for Form PTO-875

Application or Docket Number	Filing Date	☐ To be Mailed
11/338,934	01/25/2006	

APPLICATION AS FILED – PART I

FOR	(Column 1) NUMBER FILED	(Column 2) NUMBER EXTRA	SMALL ENTITY RATE ($)	SMALL ENTITY FEE ($)	OR	OTHER THAN SMALL ENTITY RATE ($)	OTHER THAN SMALL ENTITY FEE ($)
☐ BASIC FEE (37 CFR 1.16(a), (b), or (c))	N/A	N/A	N/A			N/A	
☐ SEARCH FEE (37 CFR 1.16(k), (i), or (m))	N/A	N/A	N/A			N/A	
☐ EXAMINATION FEE (37 CFR 1.16(o), (p), or (q))	N/A	N/A	N/A			N/A	
TOTAL CLAIMS (37 CFR 1.16(i))	minus 20 =	*	X $ =		OR	X $ =	
INDEPENDENT CLAIMS (37 CFR 1.16(h))	minus 3 =	*	X $ =			X $ =	
☐ APPLICATION SIZE FEE (37 CFR 1.16(s))	If the specification and drawings exceed 100 sheets of paper, the application size fee due is $250 ($125 for small entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s).						
☐ MULTIPLE DEPENDENT CLAIM PRESENT (37 CFR 1.16(j))							
* If the difference in column 1 is less than zero, enter "0" in column 2.			TOTAL			TOTAL	

APPLICATION AS AMENDED – PART II

AMENDMENT 06/02/2011	(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA	SMALL ENTITY RATE ($)	SMALL ENTITY ADDITIONAL FEE ($)	OR	OTHER THAN SMALL ENTITY RATE ($)	OTHER THAN SMALL ENTITY ADDITIONAL FEE ($)
Total (37 CFR 1.16(i))	* 11	Minus	** 20	= 0	X $ =		OR	X $52=	0
Independent (37 CFR 1.16(h))	* 6	Minus	***6	= 0	X $ =		OR	X $220=	0
☐ Application Size Fee (37 CFR 1.16(s))									
☐ FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))							OR		
					TOTAL ADD'L FEE		OR	TOTAL ADD'L FEE	0

AMENDMENT	(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA	RATE ($)	ADDITIONAL FEE ($)		RATE ($)	ADDITIONAL FEE ($)
Total (37 CFR 1.16(i))	*	Minus	**	=	X $ =		OR	X $ =	
Independent (37 CFR 1.16(h))	*	Minus	***	=	X $ =		OR	X $ =	
☐ Application Size Fee (37 CFR 1.16(s))									
☐ FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))							OR		
					TOTAL ADD'L FEE		OR	TOTAL ADD'L FEE	

* If the entry in column 1 is less than the entry in column 2, write "0" in column 3.

** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 20, enter "20".

*** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 3, enter "3".

The "Highest Number Previously Paid For" (Total or Independent) is the highest number found in the appropriate box in column 1.

Legal Instrument Examiner:
/ERIC DAVIS/

This collection of information is required by 37 CFR 1.16. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

NOTICE OF ALLOWANCE AND FEE(S) DUE

10888 7590 08/18/2011

Hollingsworth & Funk
8500 Normandale Lake Blvd., Suite 320
Minneapolis, MN 55437

EXAMINER
LEE, SIU M

ART UNIT	PAPER NUMBER
2611	

DATE MAILED: 08/18/2011

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
11/338,934	01/25/2006	Miska Hannuksela	TPO.004.C2	6610

TITLE OF INVENTION: GROUPING OF IMAGE FRAMES IN VIDEO CODING

APPLN. TYPE	SMALL ENTITY	ISSUE FEE DUE	PUBLICATION FEE DUE	PREV. PAID ISSUE FEE	TOTAL FEE(S) DUE	DATE DUE
nonprovisional	NO	$1510	$300	$0	$1810	11/18/2011

THE APPLICATION IDENTIFIED ABOVE HAS BEEN EXAMINED AND IS ALLOWED FOR ISSUANCE AS A PATENT. PROSECUTION ON THE MERITS IS CLOSED. THIS NOTICE OF ALLOWANCE IS NOT A GRANT OF PATENT RIGHTS. THIS APPLICATION IS SUBJECT TO WITHDRAWAL FROM ISSUE AT THE INITIATIVE OF THE OFFICE OR UPON PETITION BY THE APPLICANT. SEE 37 CFR 1.313 AND MPEP 1308.

THE ISSUE FEE AND PUBLICATION FEE (IF REQUIRED) MUST BE PAID WITHIN THREE MONTHS FROM THE MAILING DATE OF THIS NOTICE OR THIS APPLICATION SHALL BE REGARDED AS ABANDONED. THIS STATUTORY PERIOD CANNOT BE EXTENDED. SEE 35 U.S.C. 151. THE ISSUE FEE DUE INDICATED ABOVE DOES NOT REFLECT A CREDIT FOR ANY PREVIOUSLY PAID ISSUE FEE IN THIS APPLICATION. IF AN ISSUE FEE HAS PREVIOUSLY BEEN PAID IN THIS APPLICATION (AS SHOWN ABOVE), THE RETURN OF PART B OF THIS FORM WILL BE CONSIDERED A REQUEST TO REAPPLY THE PREVIOUSLY PAID ISSUE FEE TOWARD THE ISSUE FEE NOW DUE.

HOW TO REPLY TO THIS NOTICE:

I. Review the SMALL ENTITY status shown above.

If the SMALL ENTITY is shown as YES, verify your current SMALL ENTITY status:

A. If the status is the same, pay the TOTAL FEE(S) DUE shown above.

B. If the status above is to be removed, check box 5b on Part B - Fee(s) Transmittal and pay the PUBLICATION FEE (if required) and twice the amount of the ISSUE FEE shown above, or

If the SMALL ENTITY is shown as NO:

A. Pay TOTAL FEE(S) DUE shown above, or

B. If applicant claimed SMALL ENTITY status before, or is now claiming SMALL ENTITY status, check box 5a on Part B - Fee(s) Transmittal and pay the PUBLICATION FEE (if required) and 1/2 the ISSUE FEE shown above.

II. PART B - FEE(S) TRANSMITTAL, or its equivalent, must be completed and returned to the United States Patent and Trademark Office (USPTO) with your ISSUE FEE and PUBLICATION FEE (if required). If you are charging the fee(s) to your deposit account, section "4b" of Part B - Fee(s) Transmittal should be completed and an extra copy of the form should be submitted. If an equivalent of Part B is filed, a request to reapply a previously paid issue fee must be clearly made, and delays in processing may occur due to the difficulty in recognizing the paper as an equivalent of Part B.

III. All communications regarding this application must give the application number. Please direct all communications prior to issuance to Mail Stop ISSUE FEE unless advised to the contrary.

IMPORTANT REMINDER: Utility patents issuing on applications filed on or after Dec. 12, 1980 may require payment of maintenance fees. It is patentee's responsibility to ensure timely payment of maintenance fees when due.

PART B - FEE(S) TRANSMITTAL

Complete and send this form, together with applicable fee(s), to: Mail **Mail Stop ISSUE FEE**
Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450
or Fax (571)-273-2885

INSTRUCTIONS: This form should be used for transmitting the ISSUE FEE and PUBLICATION FEE (if required). Blocks 1 through 5 should be completed where appropriate. All further correspondence including the Patent, advance orders and notification of maintenance fees will be mailed to the current correspondence address as indicated unless corrected below or directed otherwise in Block 1, by (a) specifying a new correspondence address; and/or (b) indicating a separate "FEE ADDRESS" for maintenance fee notifications.

CURRENT CORRESPONDENCE ADDRESS (Note: Use Block 1 for any change of address)

10888 7590 08/18/2011
Hollingsworth & Funk
8500 Normandale Lake Blvd., Suite 320
Minneapolis, MN 55437

Note: A certificate of mailing can only be used for domestic mailings of the Fee(s) Transmittal. This certificate cannot be used for any other accompanying papers. Each additional paper, such as an assignment or formal drawing, must have its own certificate of mailing or transmission.

Certificate of Mailing or Transmission
I hereby certify that this Fee(s) Transmittal is being deposited with the United States Postal Service with sufficient postage for first class mail in an envelope addressed to the Mail Stop ISSUE FEE address above, or being facsimile transmitted to the USPTO (571) 273-2885, on the date indicated below.

(Depositor's name)
(Signature)
(Date)

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
11/338,934	01/25/2006	Miska Hannuksela	TPO.004.C2	6610

TITLE OF INVENTION: GROUPING OF IMAGE FRAMES IN VIDEO CODING

APPLN. TYPE	SMALL ENTITY	ISSUE FEE DUE	PUBLICATION FEE DUE	PREV. PAID ISSUE FEE	TOTAL FEE(S) DUE	DATE DUE
nonprovisional	NO	$1510	$300	$0	$1810	11/18/2011

EXAMINER	ART UNIT	CLASS-SUBCLASS
LEE, SIU M	2611	375-240120

1. Change of correspondence address or indication of "Fee Address" (37 CFR 1.363).
- ❑ Change of correspondence address (or Change of Correspondence Address form PTO/SB/122) attached.
- ❑ "Fee Address" indication (or "Fee Address" Indication form PTO/SB/47; Rev 03-02 or more recent) attached. **Use of a Customer Number is required.**

2. For printing on the patent front page, list
(1) the names of up to 3 registered patent attorneys or agents OR, alternatively,
(2) the name of a single firm (having as a member a registered attorney or agent) and the names of up to 2 registered patent attorneys or agents. If no name is listed, no name will be printed.
1 _____
2 _____
3 _____

3. ASSIGNEE NAME AND RESIDENCE DATA TO BE PRINTED ON THE PATENT (print or type)

PLEASE NOTE: Unless an assignee is identified below, no assignee data will appear on the patent. If an assignee is identified below, the document has been filed for recordation as set forth in 37 CFR 3.11. Completion of this form is NOT a substitute for filing an assignment.

(A) NAME OF ASSIGNEE (B) RESIDENCE: (CITY and STATE OR COUNTRY)

Please check the appropriate assignee category or categories (will not be printed on the patent) : ❑ Individual ❑ Corporation or other private group entity ❑ Government

4a. The following fee(s) are submitted:
- ❑ Issue Fee
- ❑ Publication Fee (No small entity discount permitted)
- ❑ Advance Order - # of Copies _____

4b. Payment of Fee(s): **(Please first reapply any previously paid issue fee shown above)**
- ❑ A check is enclosed.
- ❑ Payment by credit card. Form PTO-2038 is attached.
- ❑ The Director is hereby authorized to charge the required fee(s), any deficiency, or credit any overpayment, to Deposit Account Number _____ (enclose an extra copy of this form).

5. **Change in Entity Status** (from status indicated above)
- ❑ a. Applicant claims SMALL ENTITY status. See 37 CFR 1.27.
- ❑ b. Applicant is no longer claiming SMALL ENTITY status. See 37 CFR 1.27(g)(2).

NOTE: The Issue Fee and Publication Fee (if required) will not be accepted from anyone other than the applicant; a registered attorney or agent; or the assignee or other party in interest as shown by the records of the United States Patent and Trademark Office.

Authorized Signature _____ Date _____

Typed or printed name _____ Registration No. _____

This collection of information is required by 37 CFR 1.311. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, Virginia 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, Virginia 22313-1450.

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
11/338,934	01/25/2006	Miska Hannuksela	TPO.004.C2	6610

10888 7590 08/18/2011

Hollingsworth & Funk
8500 Normandale Lake Blvd., Suite 320
Minneapolis, MN 55437

EXAMINER	
LEE, SIU M	

ART UNIT	PAPER NUMBER
2611	

DATE MAILED: 08/18/2011

Determination of Patent Term Adjustment under 35 U.S.C. 154 (b)
(application filed on or after May 29, 2000)

The Patent Term Adjustment to date is 1107 day(s). If the issue fee is paid on the date that is three months after the mailing date of this notice and the patent issues on the Tuesday before the date that is 28 weeks (six and a half months) after the mailing date of this notice, the Patent Term Adjustment will be 1107 day(s).

If a Continued Prosecution Application (CPA) was filed in the above-identified application, the filing date that determines Patent Term Adjustment is the filing date of the most recent CPA.

Applicant will be able to obtain more detailed information by accessing the Patent Application Information Retrieval (PAIR) WEB site (http://pair.uspto.gov).

Any questions regarding the Patent Term Extension or Adjustment determination should be directed to the Office of Patent Legal Administration at (571)-272-7702. Questions relating to issue and publication fee payments should be directed to the Customer Service Center of the Office of Patent Publication at 1-(888)-786-0101 or (571)-272-4200.

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether disclosure of these records is required by the Freedom of Information Act.
2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.
3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.
4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).
5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.
6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).
7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.
8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspection or an issued patent.
9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

Notice of Allowability	Application No.	Applicant(s)
	11/338,934	HANNUKSELA, MISKA
	Examiner	Art Unit
	SIU LEE	2611

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address--

All claims being allowable, PROSECUTION ON THE MERITS IS (OR REMAINS) CLOSED in this application. If not included herewith (or previously mailed), a Notice of Allowance (PTOL-85) or other appropriate communication will be mailed in due course. **THIS NOTICE OF ALLOWABILITY IS NOT A GRANT OF PATENT RIGHTS.** This application is subject to withdrawal from issue at the initiative of the Office or upon petition by the applicant. See 37 CFR 1.313 and MPEP 1308.

1. ☒ This communication is responsive to *amendment filed on 6/2/2011*.
2. ☒ The allowed claim(s) is/are *1,2,4-6,8-11,13 and 14*.
3. ☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
   a) ☒All b) ☐ Some* c) ☐ None of the:
      1. ☒ Certified copies of the priority documents have been received.
      2. ☐ Certified copies of the priority documents have been received in Application No. _____ .
      3. ☐ Copies of the certified copies of the priority documents have been received in this national stage application from the International Bureau (PCT Rule 17.2(a)).

   * Certified copies not received: _____.

Applicant has THREE MONTHS FROM THE "MAILING DATE" of this communication to file a reply complying with the requirements noted below. Failure to timely comply will result in ABANDONMENT of this application.
THIS THREE-MONTH PERIOD IS NOT EXTENDABLE.

4. ☐ A SUBSTITUTE OATH OR DECLARATION must be submitted. Note the attached EXAMINER'S AMENDMENT or NOTICE OF INFORMAL PATENT APPLICATION (PTO-152) which gives reason(s) why the oath or declaration is deficient.
5. ☐ CORRECTED DRAWINGS (as "replacement sheets") must be submitted.
   (a) ☐ including changes required by the Notice of Draftsperson's Patent Drawing Review (PTO-948) attached
      1) ☐ hereto or 2) ☐ to Paper No./Mail Date _____.
   (b) ☐ including changes required by the attached Examiner's Amendment / Comment or in the Office action of Paper No./Mail Date _____.

   Identifying indicia such as the application number (see 37 CFR 1.84(c)) should be written on the drawings in the front (not the back) of each sheet. Replacement sheet(s) should be labeled as such in the header according to 37 CFR 1.121(d).
6. ☐ DEPOSIT OF and/or INFORMATION about the deposit of BIOLOGICAL MATERIAL must be submitted. Note the attached Examiner's comment regarding REQUIREMENT FOR THE DEPOSIT OF BIOLOGICAL MATERIAL.

Attachment(s)

1. ☐ Notice of References Cited (PTO-892)	5. ☐ Notice of Informal Patent Application
2. ☐ Notice of Draftperson's Patent Drawing Review (PTO-948)	6. ☐ Interview Summary (PTO-413), Paper No./Mail Date _____ .
3. ☒ Information Disclosure Statements (PTO/SB/08), Paper No./Mail Date 4/5/2011, 6/2/2011	7. ☐ Examiner's Amendment/Comment
4. ☐ Examiner's Comment Regarding Requirement for Deposit of Biological Material	8. ☒ Examiner's Statement of Reasons for Allowance
	9. ☐ Other _____.

DETAILED ACTION

Allowable Subject Matter

1. Claims 1-2, 4-6, 8-11, and 13-14 are allowed.

2. The following is an examiner's statement of reasons for allowance:

(1) Regarding claims 1-2, 4-6, and 8-9:

The present invention describes a method, encoder and a computer program product for encoding a video sequence comprising an independent sequence of image frames, wherein all motion-compensated temporal prediction references of the independent sequence refer only to image frames within said independent sequence, the method comprising encoding into the video sequence an indication of at least one image frame, which is the first image frame, in decoding order, of the independent sequence; encoding identifier values for the image frames according to a numbering scheme; and resetting the identifier value for the indicated first image frame of the independent sequence. The closest prior art, Apostolopoulos (US 7,103,669 B2) discloses a similar method but fails to disclose encoding identifier values for the image frames according to a numbering scheme; and resetting the identifier value for the indicated first image frame of the independent sequence. This distinct feature has been added to the independent claims 1, 5, and 9, thus rendering claims 1-2, 4-6, and 8-9 allowable.

(2) Regarding claims 10-11, 13-14:

The present invention describes a method, decoder, and a computer program product for decoding a compressed video sequence, comprising decoding from the video sequence an indication of at least one image frame, which is the first image frame, in decoding order, of an independent sequence, wherein all motion-compensated temporal prediction references of the independent sequence refer only to image frames within said independent sequence; starting the decoding of the video sequence from said first image frame of the independent sequence, whereby the video sequence is decoded without prediction from any image frame decoded prior to said first image frame; decoding identifier values for image frames according to a numbering scheme; and resetting the identifier value for the indicated first image frame of the independent sequence. The closest prior art, Apostolopoulos (US 7,103,669 B2) discloses a similar method but fails to disclose starting the decoding of the video sequence from said first image frame of the independent sequence, whereby the video sequence is decoded without prediction from any image frame decoded prior to said first image frame; decoding identifier values for image frames according to a numbering scheme; and resetting the identifier value for the indicated first image frame of the independent sequence. This distinct feature has been added to independent claims 10, 13 and 14, thus rendering claims 10-11 and 13-14 allowable.

Any comments considered necessary by applicant must be submitted no later than the payment of the issue fee and, to avoid processing delays, should preferably

accompany the issue fee. Such submissions should be clearly labeled "Comments on Statement of Reasons for Allowance."

Any inquiry concerning this communication or earlier communications from the examiner should be directed to SIU LEE whose telephone number is (571)270-1083. The examiner can normally be reached on Mon-Fri, 6:15-3:45 with every other Friday off.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Chieh Fan can be reached on (571) 272-3042. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Siu M Lee/
Examiner, Art Unit 2611
8/12/2011
/Shuwang Liu/
Supervisory Patent Examiner, Art Unit 2611

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

BIB DATA SHEET

CONFIRMATION NO. 6610

SERIAL NUMBER	FILING or 371(c) DATE RULE	CLASS	GROUP ART UNIT	ATTORNEY DOCKET NO.
11/338,934	01/25/2006	375	2611	TPO.004.C2

APPLICANTS
Miska Hannuksela, Tampere, FINLAND;

**** CONTINUING DATA ***************************
This application is a CON of 10/306,942 11/29/2002 PAT 7,894,521

**** FOREIGN APPLICATIONS ***************************
FINLAND 20020127 01/23/2002

**** IF REQUIRED, FOREIGN FILING LICENSE GRANTED ****
02/28/2006

Foreign Priority claimed ☑ Yes ☐ No 35 USC 119(a-d) conditions met ☑ Yes ☐ No Verified and Acknowledged /SIU M LEE/ Examiner's Signature	☐ Met after Allowance Initials	STATE OR COUNTRY	SHEETS DRAWINGS	TOTAL CLAIMS	INDEPENDENT CLAIMS
		FINLAND	4	16	7

ADDRESS
Hollingsworth & Funk
8500 Normandale Lake Blvd., Suite 320
Minneapolis, MN 55437
UNITED STATES

TITLE
Grouping of image frames in video coding

FILING FEE RECEIVED		
1800	FEES: Authority has been given in Paper No._____ to charge/credit DEPOSIT ACCOUNT No._____ for following:	☐ All Fees ☐ 1.16 Fees (Filing) ☐ 1.17 Fees (Processing Ext. of time) ☐ 1.18 Fees (Issue) ☐ Other _____ ☐ Credit

Search Notes	Application/Control No.	Applicant(s)/Patent Under Reexamination
	11338934	HANNUKSELA, MISKA
	Examiner	Art Unit
	SIU M LEE	2611

SEARCHED			
Class	Subclass	Date	Examiner
375	240.12, 240.25	5/21/2010	Siu M. Lee

SEARCH NOTES		
Search Notes	Date	Examiner
EAST	5/21/2010	Siu M. Lee
Inventor name search - Miska hannuksela	5/21/2010	Siu M. Lee
Discuss the application with Chieh Fan	5/21/2010	Siu M. Lee
EAST	12/1/2010	Siu M. Lee
Discuss the amendment with Chieh Fan	12/1/2010	Siu M. Lee
Discuss the 101 and 112 rejection of claims 5-8 and 13 with John Peng	12/1/2010	Siu M. Lee
EAST	2/28/2011	Siu M. Lee
Discuss the amendment with regard to 101 problem with Daniel Swerdlow	2/25/2011	Siu M. Lee
EAST	8/12/2011	Siu M. Lee
Discussed the amendment with Chieh Fan	8/8/2011	Siu M. Lee
Discussed the amendment with Shuwang Liu	8/9/2011	Siu M. Lee
Discussed the application with WenPeng Chen	8/11/2011	Siu M. Lee

INTERFERENCE SEARCH			
Class	Subclass	Date	Examiner
375	240.12, 240.25	8/12/2011	Siu M. Lee

FORM 1449* SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT IN AN APPLICATION (Use several sheets if necessary)	Docket Number: TPO.004.C2	Application Number: 11/338,934
	Applicant: HANNUKSELA	
	Filing Date: 01/25/2006	Group Art Unit: 2611

U.S. PATENT DOCUMENTS

EXAMINER INITIAL	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE

FOREIGN PATENT DOCUMENTS

	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	TRANSLATION	
						YES	NO

OTHER DOCUMENTS (Including Author, Title, Date, Pertinent Pages, Etc.)

/SL/	2002	Hannuksela, "Signaling of Enhanced GOPs", Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, 4th Meeting, Austria, July 22-26, 2002.
/SL/	4/5/2011	File History for EP Application No. 03700321.7 as retrieved from the European Patent Office Electronic File System on 4/5/2011, 222 pages.

EXAMINER /Siu Lee/	DATE CONSIDERED 08/12/2011

EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609; draw line through citation if not in conformance and not considered. Include copy of this form for next communication to the Applicant.

*Substitute Disclosure Statement Form (PTO-1449) Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

EAST Search History

EAST Search History (Prior Art)

Ref #	Hits	Search Query	DBs	Default Operator	Plurals	Time Stamp
S1	0	11/338934 and i(dentifier with independent sequence)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2011/08/04 14:53
S2	1	11/338934 and (identifier with independent sequence)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2011/08/04 14:54
S3	14	identifier with (gop or group near2 picture or independent sequence) with encod$3	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2011/08/04 14:59
S4	1654	MPEG and (reset$4 with frame)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2011/08/05 13:53
S5	1653	S4 and (@pd<="20020123" OR @ad<="200230123" OR @rlad<="20020123" OR @prad<="20020123")	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2011/08/05 13:54
S6	27	S5 and (MPEG with(reset$4 with frame))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2011/08/05 13:54
S7	3	MPEG same2(I frame with (reset$4 with gop))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2011/08/05 14:17
S8	13	MPEG same2 ((intra or I) with (reset$4 with gop))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2011/08/05 14:18
S9	13	S8 and (@pd<="20020123" OR @ad<="200230123" OR @rlad<="20020123" OR @prad<="20020123")	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2011/08/05 14:18
S10	7	S8 and (@pd<="20020123" OR @ad<="20020123" OR @rlad<="20020123" OR @prad<="20020123")	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2011/08/05 14:21

S11	8	(group of picture or gop) with reset$4 with I frame	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2011/08/07 16:13
S12	48	MPEG same (Gop with reset$4)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2011/08/07 16:51
S13	48	S12 and (@pd<="20020123" OR @ad<="200230123" OR @rlad<="20020123" OR @prad<="20020123")	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2011/08/07 16:51
S14	31	S12 and (@pd<="20020123" OR @ad<="20020123" OR @rlad<="20020123" OR @prad<="20020123")	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2011/08/07 16:58
S15	0	S13 and transcoder	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2011/08/07 16:59
S16	6	S13 and transcod$3	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2011/08/07 16:59
S17	782	(random$2 near2 access) with frame with (decod$3 or decompress$3)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2011/08/08 10:07
S18	212123	(mpeg or video) with (encod$3 or cod$3 or decod$3 or decompress$3)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2011/08/08 10:08
S19	301	S17 same2 S18	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2011/08/08 10:08
S20	123	S19 and header	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2011/08/08 10:08
S22	53	S20 and (@pd<="20020123" OR @ad<="20020123" OR @rlad<="20020123" OR @prad<="20020123")	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2011/08/08 10:09
S23	1	(reset$4 with index with (frame or image)) and (gop or group of picture)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2011/08/08 10:18

S24	71	(reset$4 with index) and (gop or group of picture)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2011/08/08 10:19
S25	42	(reset$4 with index with (frame or image)) and (gop or group of picture or mpeg)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2011/08/08 10:19
S26	10	S25 and (@pd<="20020123" OR @ad<="20020123" OR @rlad<="20020123" OR @prad<="20020123")	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2011/08/08 10:19
S27	171	(reset$4 near3 inde$3) and (gop or group of picture or mpeg)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2011/08/08 10:20
S28	99	S27 and (@pd<="20020123" OR @ad<="20020123" OR @rlad<="20020123" OR @prad<="20020123")	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2011/08/08 10:20
S29	44	S28 and ((reset$4 near3 inde$3 with (frame or picture or image)) and (gop or group of picture or mpeg))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2011/08/08 10:21
S30	41	S28 and ((reset$4 near3 inde$3 near3 (frame or picture or image)) and (gop or group of picture or mpeg))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2011/08/08 10:22
S32	2	(scene near2 change) with reset$4 with I frame	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2011/08/08 10:27
S33	1	11/338934 and reset$4	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2011/08/08 10:29
S34	0	reset$4 with (intra frame) with S17	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2011/08/08 10:34
S35	0	reset$4 with (intra frame) with "0"	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2011/08/08 10:34
S36	30	reset$4 with (I frame) with "0"	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2011/08/08 10:35

S37	31	reset$4 with (I frame) with ("0" or S17)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2011/08/08 10:35
S38	1	"20030039308".pn. and frame counter	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2011/08/08 11:18

EAST Search History (Interference)

Ref #	Hits	Search Query	DBs	Default Operator	Plurals	Time Stamp
L1	5528	375/240.12,240.25.ccls.	US-PGPUB; USPAT; UPAD	ADJ	ON	2011/08/12 10:21
L2	2334	1 and (@pd<="20020123" OR @ad<="20020123" OR @rlad<="20020123" OR @prad<="20020123")	US-PGPUB; USPAT; UPAD	ADJ	ON	2011/08/12 10:26
L3	166	2 and ((encod$3 or cod$3 or decod$3) with (indentifier or number$3 or indicat$3) with (image or video) frame)	US-PGPUB; USPAT; UPAD	ADJ	ON	2011/08/12 10:56
L4	30	3 and reset$4	US-PGPUB; USPAT; UPAD	ADJ	ON	2011/08/12 10:56
L5	3755	((encod$3 or cod$3 or decod$3) with (indentifier or number$3 or indicat$3) with (image or video) frame)	US-PGPUB; USPAT; UPAD	ADJ	ON	2011/08/12 11:29
L7	2178	5 and (@pd<="20020123" OR @ad<="20020123" OR @rlad<="20020123" OR @prad<="20020123")	US-PGPUB; USPAT; UPAD	ADJ	ON	2011/08/12 11:30
L8	390	7 and reset$4	US-PGPUB; USPAT; UPAD	ADJ	ON	2011/08/12 11:30
L9	19	7 and (reset$4 with (indentifier or number$3 or indicat$3) with (image or video) frame)	US-PGPUB; USPAT; UPAD	ADJ	ON	2011/08/12 11:31
L10	15	9 not 4	US-PGPUB; USPAT; UPAD	ADJ	ON	2011/08/12 11:31
L11	2030	((encod$3 or cod$3 or decod$3) with (reference) with (image or video) frame)	US-PGPUB; USPAT; UPAD	ADJ	ON	2011/08/12 11:38
L12	937	11 and (@pd<="20020123" OR @ad<="20020123" OR @rlad<="20020123" OR @prad<="20020123")	US-PGPUB; USPAT; UPAD	ADJ	ON	2011/08/12 11:38

L13	3	12 and (reset$4 with (reference) with (image or video) frame)	US-PGPUB; USPAT; UPAD	ADJ	ON	2011/08/12 11:39

8/ 12/ 2011 11:41:23 AM
C:\ Users\ slee7\ Documents\ EAST\ Workspaces\ 11338934 201107.wsp

Index of Claims	Application/Control No.	Applicant(s)/Patent Under Reexamination
	11338934	HANNUKSELA, MISKA
	Examiner	Art Unit
	SIU M LEE	2611

✓	Rejected	-	Cancelled	N	Non-Elected	A	Appeal
=	Allowed	÷	Restricted	I	Interference	O	Objected

☐ Claims renumbered in the same order as presented by applicant ☐ CPA ☐ T.D. ☐ R.1.47

CLAIM		DATE								
Final	Original	05/21/2010	12/02/2010	02/28/2011	08/12/2011					
1	1	✓	O	✓	=					
2	2	✓	O	✓	=					
	3	✓	-	-	-					
3	4	✓	O	✓	=					
4	5	✓	✓	✓	=					
5	6	✓	✓	✓	=					
	7	✓	-	-	-					
6	8	✓	✓	✓	=					
7	9	✓	O	✓	=					
8	10	✓	O	✓	=					
9	11	✓	O	✓	=					
	12	✓	-	-	-					
10	13	✓	✓	✓	=					
11	14	✓	O	✓	=					
	15	✓	-	-	-					
	16	✓	-	-	-					

FORM 1449* INFORMATION DISCLOSURE STATEMENT IN AN APPLICATION (Use several sheets if necessary)	Docket Number: TPO.004.C2	Application Number: 11/338,934
	Applicant: HANNUKSELA	
	Filing Date: 1/25/2006	Group Art Unit: 2611

U.S. PATENT DOCUMENTS

EXAMINER INITIAL	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE

FOREIGN PATENT DOCUMENTS

	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	TRANSLATION	
						YES	NO

OTHER DOCUMENTS (Including Author, Title, Date, Pertinent Pages, Etc.)

/SL/	2/18/2011	Office Action dated 2/18/2011 from U.S. Serial No. 11/338,996, 18 pages.
/SL/	5/17/2011	Office Action Response dated 5/17/2011 from U.S. Serial No. 11/338,996, 9 pages.

EXAMINER /Siu Lee/	DATE CONSIDERED 08/12/2011

EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609; draw line through citation if not in conformance and not considered. Include copy of this form for next communication to the Applicant.

*Substitute Disclosure Statement Form (PTO-1449) Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Issue Classification	Application/Control No.	Applicant(s)/Patent Under Reexamination
	11338934	HANNUKSELA, MISKA
	Examiner	**Art Unit**
	SIU LEE	2611

ORIGINAL		INTERNATIONAL CLASSIFICATION														
CLASS	SUBCLASS	CLAIMED									NON-CLAIMED					
375	240.12	H	0	4	N	7 / 12 (2006.0)										

CROSS REFERENCE(S)					
CLASS	SUBCLASS (ONE SUBCLASS PER BLOCK)				
375	240.12	240.25			

☐ Claims renumbered in the same order as presented by applicant ☐ CPA ☐ T.D. ☐ R.1.47

Final	Original	Final	Original	Final	Original	Final	Original	Final	Original	Final	Original	Final	Original	Final	Original
1	1														
2	2														
	3														
3	4														
4	5														
5	6														
	7														
6	8														
7	9														
8	10														
9	11														
	12														
10	13														
11	14														
	15														
	16														

/SIU LEE/ Examiner.Art Unit 2611	8/12/2011	**Total Claims Allowed:**	
(Assistant Examiner)	(Date)	11	
/SHUWANG LIU/ Supervisory Patent Examiner.Art Unit 2611	08/12/2011	O.G. Print Claim(s)	O.G. Print Figure
(Primary Examiner)	(Date)	1	4

FORM 1449* INFORMATION DISCLOSURE STATEMENT IN AN APPLICATION (Use several sheets if necessary)	Docket Number: TPO.004.C2	Application Number: 11/338,934
	Applicant: HANNUKSELA	
	Filing Date: 1/25/2006	Group Art Unit: 2611

U.S. PATENT DOCUMENTS

EXAMINER INITIAL	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE

FOREIGN PATENT DOCUMENTS

	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	TRANSLATION	
						YES	NO
	JP 8265694	10/11/1996	JP			Abstract	

OTHER DOCUMENTS (Including Author, Title, Date, Pertinent Pages, Etc.)

	8/2/2011	Notice of Allowance dated 8/2/2011 from Japanese Application No. 2006-120296, 3 pages.

EXAMINER	DATE CONSIDERED
EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609; draw line through citation if not in conformance and not considered. Include copy of this form for next communication to the Applicant.	

*Substitute Disclosure Statement Form (PTO-1449) Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Electronic Patent Application Fee Transmittal

Application Number:	11338934
Filing Date:	25-Jan-2006
Title of Invention:	GROUPING OF IMAGE FRAMES IN VIDEO CODING
First Named Inventor/Applicant Name:	Miska Hannuksela
Filer:	Erin M. Nichols
Attorney Docket Number:	TPO.004.C2

Filed as Large Entity

Utility under 35 USC 111(a) Filing Fees

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Basic Filing:				
Pages:				
Claims:				
Miscellaneous-Filing:				
Petition:				
Patent-Appeals-and-Interference:				
Post-Allowance-and-Post-Issuance:				
Extension-of-Time:				

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Miscellaneous:				
Submission- Information Disclosure Stmt	1806	1	180	180
		Total in USD ($)		**180**

Electronic Acknowledgement Receipt

EFS ID:	10746513
Application Number:	11338934
International Application Number:	
Confirmation Number:	6610
Title of Invention:	GROUPING OF IMAGE FRAMES IN VIDEO CODING
First Named Inventor/Applicant Name:	Miska Hannuksela
Customer Number:	10888
Filer:	Erin M. Nichols
Filer Authorized By:	
Attorney Docket Number:	TPO.004.C2
Receipt Date:	19-AUG-2011
Filing Date:	25-JAN-2006
Time Stamp:	15:33:33
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	yes
Payment Type	Deposit Account
Payment was successfully received in RAM	$180
RAM confirmation Number	1793
Deposit Account	503581
Authorized User	
The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows: Charge any Additional Fees required under 37 C.F.R. Section 1.17 (Patent application and reexamination processing fees)	

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Foreign Reference	JP8265694.pdf	1970587 9b6b0c01709297dce8c4821c1d134fe43da4ece0	no	22
Warnings:					
Information:					
2	Non Patent Literature	JPAllowance.pdf	51955 b2d041bac2aacea9ef9b9171477eb66f75eb9aea	no	3
Warnings:					
Information:					
3	Transmittal Letter	IDS-15-TPO004C2.pdf	56181 565c9d2bcf721f3d5a30a6ce6c6ab5f37bc851e0	no	2
Warnings:					
Information:					
4	Information Disclosure Statement (IDS) Form (SB08)	1449-15-TPO004C2.pdf	54443 440f5de27d3aa13d49423baa07fc3da01bf8a8cb	no	1
Warnings:					
Information:					
This is not an USPTO supplied IDS fillable form					
5	Fee Worksheet (SB06)	fee-info.pdf	30352 7e704e38f19748d26523393c73c2d73bc43627af	no	2
Warnings:					
Information:					
		Total Files Size (in bytes):	2163518		

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111

If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371

If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office

If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

Espacenet

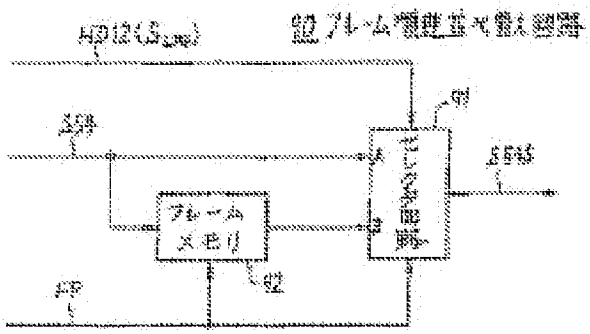
Bibliographic data: JP 8265694 (A)

ENCODED DATA DECODING METHOD AND DEVICE

Publication date:	1996-10-11
Inventor(s):	YONEMITSU JUN +
Applicant(s):	SONY CORP +
Classification:	**- international:** ***H04N5/91; H04N5/92; H04N7/24; H04N7/26;*** (IPC1-7): H04N5/91; H04N5/92; H04N7/24 **- European:**
Application number:	JP19950352516 19951230
Priority number (s):	JP19950352516 19951230
Also published as:	• JP 2874745 (B2)

Abstract of JP 8265694 (A)

PURPOSE: To rightly restore the encoded image data after an edition is performed by preventing the image for which a second inter-encoding is performed from being outputted by using the image belonging to an original preceding group as a predicted image when an edition flag is set. CONSTITUTION: By using the image belonging to an original preceding group as a predicted image, an edition flag SLPG is set to the group including the encoded data corresponding to the image for which a second inter-encoding which is possible to use the images existing before and after in the time in a display order as predicted images is performed.; A frame order reverse rearranging circuit 90 executes the processing rearranging from the frame order according to a decoding processing to the frame order based on an inut image signal and executes an edition reproduction processing by referring to the edition flag SLPG. Namely, when the edition flag SLPG is set, the image belonging to the original preceding frame group is used as the the predicted image, an interpolation frame is not outputted, the intraframe encoded just before is outputted and an encoding is performed.



Last updated: 26.04.2011 Worldwide Database 5.7.23; 93p

(19)日本国特許庁（ＪＰ）　(12) **公開特許公報**（Ａ）　(11)特許出願公開番号

特開平8－265694

(43)公開日　平成８年(1996)10月11日

(51)Int.Cl.[6]		識別記号	庁内整理番号	ＦＩ			技術表示箇所
Ｈ０４Ｎ	5/92			Ｈ０４Ｎ	5/92	Ｈ	
	5/91				5/91	Ｎ	
	7/24				7/13	Ｚ	

審査請求　有　　請求項の数４　ＦＤ　（全 21 頁）

(21)出願番号　特願平7－352516
(62)分割の表示　特願平2－119604の分割
(22)出願日　平成２年(1990) ５月９日

(71)出願人　000002185
ソニー株式会社
東京都品川区北品川６丁目７番35号
(72)発明者　米満　潤
東京都品川区北品川６丁目７番35号ソニー株式会社内
(74)代理人　弁理士　田辺　恵基

(54)【発明の名称】　符号化データ復号化方法及び符号化データ復号化装置

(57)【要約】
【課題】画像信号を複数のフレームを含む複数のグループに分割して符号化した場合、このグループを分断するような編集を行うと、編集後の画像を再生した際の画質が劣化する問題があつた。
【解決手段】編集フラグの有無を検出し、編集フラグが設定されていた場合には先行するグループに属する画像を予測画像に用いて、表示順上で時間的に前後にある画像を予測画像として用いることが可能なインター符号化によつて生成された画像を出力しないようにしたことにより、編集後の符号化画像データを正しく復元し得る符号化データ復号化方法及び装置を実現できる。

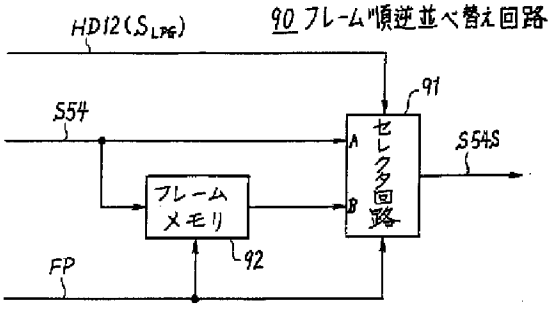


図１２　フレーム順逆並べ替え回路

【特許請求の範囲】
【請求項１】イントラ符号化によつて生成された第１の符号化データと、表示順上で時間的に前にある画像のみを予測画像として用いることが可能な第１のインター符号化によつて生成された第２の符号化データと、表示順上で時間的に前後にある画像を予測画像として用いることが可能な第２のインター符号化によつて生成された第３の符号化データとを含む符号化画像データが、それぞれいくつかの画像に対応する符号化データを含む複数のグループにグループ化され、上記符号化画像データは、元の先行するグループとは異なるグループに先行されるあるグループを少なくとも１つ含むように編集されてなり、上記あるグループは、上記元の先行するグループに属する画像を上記予測画像に用いて上記第２のインター符号化がなされた画像に対応する符号化データを含み、上記あるグループには編集フラグが設定されているような符号化画像データを復号化する符号化データ復号化方法において、
上記編集フラグの設定の有無を検出し、
上記編集フラグが設定されていた場合、上記元の先行するグループに属する画像を上記予測画像に用いて第２のインター符号化された画像を出力しないようにすることを特徴とする符号化データ復号化方法。
【請求項２】上記元の先行するグループに属する画像を上記予測画像に用いて第２のインター符号化された画像のかわりに、直前に復号化されたイントラ符号化された画像を出力することを特徴とする請求項１に記載の符号化データ復号化方法。
【請求項３】イントラ符号化によつて生成された第１の符号化データと、表示順上で時間的に前にある画像のみを予測画像として用いることが可能な第１のインター符号化によつて生成された第２の符号化データと、表示順上で時間的に前後にある画像を予測画像として用いることが可能な第２のインター符号化によつて生成された第３の符号化データとを含む符号化画像データが、それぞれいくつかの画像に対応する符号化データを含む複数のグループにグループ化され、上記符号化画像データは、元の先行するグループとは異なるグループに先行されるあるグループを少なくとも１つ含むように編集されてなり、上記あるグループは、上記元の先行するグループに属する画像を上記予測画像に用いて上記第２のインター符号化がなされた画像に対応する符号化データを含み、上記あるグループには編集フラグが設定されているような符号化画像データを復号化する符号化データ復号化装置において、
上記編集フラグの設定の有無を検出する手段と、
上記編集フラグが設定されていた場合、上記元の先行するグループに属する画像を上記予測画像に用いて第２のインター符号化された画像を出力しないようにする復号化手段とを具えることを特徴とする符号化データ復号化装置。
【請求項４】上記復号化手段は、上記元の先行するグループに属する画像を上記予測画像に用いて第２のインター符号化された画像のかわりに、直前に復号化されたイントラ符号化された画像を出力することを特徴とする請求項３に記載の符号化データ復号化装置。
【発明の詳細な説明】
【０００１】
【目次】以下の順序で本発明を説明する。
発明の属する技術分野
従来の技術（図１６～図１８）
発明が解決しようとする課題（図１９）
課題を解決するための手段
発明の実施の形態
（１）動画符号化／復号化装置の全体構成（図１～図５）
（２）実施例による動画符号化データの記録順序（図６～図９）
（３）実施例による動画符号化データのエデイツト処理（図６～図１１）
（４）実施例によるエデイツトされた符号化データの復号化方法（図８及び図１２）
（５）他の実施例（図１３～図１５）
発明の効果
【０００２】
【発明の属する技術分野】本発明は符号化データ復号化方法及び装置に関し、例えばエデイツト可能に符号化した符号化データを復号する際に適用して好適なものである。
【０００３】
【従来の技術】従来動画映像でなる映像信号をフレーム内符号化データ及びフレーム間符号化データに高能率符号化して、例えば光磁気デイスク構成のコンパクトデイスク（ＣＤ－ＭＯデイスク）に高密度記録し、当該記録された動画符号化データを必要に応じて検索して再生し得るようになされた記録再生装置が提案されている（特開昭63-1183号公報、特願平1-267049号）。
【０００４】すなわち、例えば図１６（Ａ）に示すように、時点ｔ＝t_1、t_2、t_3……において動画を構成する各画像ＰＣ１、ＰＣ２、ＰＣ３……をデイジタル符号化して、例えばＣＤ－ＭＯ記録再生装置でなる伝送系に伝送する場合、映像信号には時間の経過に従つて自己相関が大きい特徴がある点を利用して伝送すべき画像データを圧縮処理することにより伝送効率を高めるような処理をするもので、フレーム内符号化処理は画像ＰＣ１、ＰＣ２、ＰＣ３……を例えば画素データを所定の基準値と比較して差分を求めるような圧縮処理を実行し、かくして各画像ＰＣ１、ＰＣ２、ＰＣ３……について同一フレーム内における画素データ間の自己相関を利用して圧縮されたデータ量の画像データを伝送する。

【０００５】またフレーム間符号化処理は、図１６（Ｂ）に示すように、順次隣合う画像ＰＣ１及びＰＣ２、ＰＣ２及びＰＣ３……間の画素データの差分でなる画像データＰＣ１２、ＰＣ２３……を求め、これを時点ｔ＝t_1における初期画像ＰＣ１についてフレーム内符号化処理された画像データと共に伝送する。かくして画像ＰＣ１、ＰＣ２、ＰＣ３……をそのすべての画像データを伝送する場合と比較して格段的にデータ量が少ないデイジタルデータに高能率符号化してＣＤ－ＭＯ記録再生装置に伝送することができる。

【０００６】かかる映像信号の符号化処理は、図１７に示す構成の動画符号化データ発生装置１において実行される。動画符号化データ発生装置１は入力映像信号ＶＤを前処理回路２において処理することにより片フイールド落し処理及び片フイールドライン間引き処理等の処理をした後、輝度信号及びクロマ信号を16画素（水平方向に）×16画素（垂直方向に）分のデータでなる伝送単位ブロツク（これをマクロブロツクと呼ぶ）データＳ１１に変換して画像データ符号化回路３に供給する。

【０００７】画像データ符号化回路３は予測符号化回路４において形成される予測現フレームデータＳ１２を受けてマクロブロツクデータＳ１１との差分を求めることによつてフレーム間符号化データを発生し（これをフレーム間符号化モードと呼ぶ）、又はマクロブロツクデータＳ１１と基準値データとの差分を求めることによりフレーム内符号化データを形成してこれを差分データＳ１３として変換符号化回路５に供給する。

【０００８】変換符号化回路５はデイスクリートコサイン変換回路で構成され、差分データＳ１３を直交変換することによつて高能率符号化してなる変換符号化データＳ１４を量子化回路６に与えることにより量子化画像データＳ１５を送出させる。かくして量子化回路６から得られる量子化画像データＳ１５は可変長符号化回路を含んでなる再変換符号化回路７において再度高能率符号化処理された後、伝送画像データＳ１６として伝送バツフアメモリ８に供給される。

【０００９】これに加えて量子化画像データＳ１５は予測符号化回路４において逆量子化、逆変換符号化処理されることより差分データに復号化された後予測前フレームデータを差分データによつて修正演算することにより新たな予測前フレームデータを保存すると共に、マクロブロツクデータＳ１１に基づいて形成される動き検出データによつて予測符号化回路４に保存されている予測前フレームデータを動き補償することにより予測現フレームデータを形成して画像データ符号化回路３に供給できるようになされ、これにより現在伝送しようとするフレーム（すなわち現フレーム）のマクロブロツクデータＳ１１と予測現フレームデータＳ１２との差分を差分データＳ１３として得るようになされている。

【００１０】図１７の構成において、図１６について上述した動画像を伝送する場合、先ず図１６（Ａ）の時点t_1において画像ＰＣ１の画像データがマクロブロツクデータＳ１１として与えられたとき、画像データ符号化回路３はフレーム内符号化モードになつてこれをフレーム内符号化処理された差分データＳ１３として変換符号化回路５に供給し、これにより量子化回路６、再変換符号化回路７を介して伝送バツフアメモリ８に伝送画像データＳ１６を供給する。

【００１１】これと共に、量子化回路６の出力端に得られる量子化画像データＳ１５が予測符号化回路４において予測符号化処理されることにより、伝送バツフアメモリ８に送出された伝送画像データＳ１６を表す予測前フレームデータが前フレームメモリに保持され、続いて時点t_2において画像ＰＣ２を表すマクロブロツクデータＳ１１が画像データ符号化回路３に供給されたとき、予測現フレームデータＳ１２に動き補償されて画像データ符号化回路３に供給される。

【００１２】かくして時点ｔ＝t_2において画像データ符号化回路３はフレーム間符号化処理された差分データＳ１３を変換符号化回路５に供給し、これにより当該フレーム間の画像の変化を表す差分データが伝送画像データＳ１６として伝送バツフアメモリ８に供給されると共に、その量子化画像データＳ１５が予測符号化回路４に供給されることにより予測符号化回路４において予測前フレームデータが形成されると共に保存される。

【００１３】以下同様の動作が繰り返されることにより、画像データ符号化回路３がフレーム間符号化処理を実行している間、前フレームと現フレームとの間の画像の変化を表す差分データだけが伝送バツフアメモリ８に順次送出されることになる。伝送バツフアメモリ８はこのようにして送出されて来る伝送画像データＳ１６を一旦記憶し、伝送路９の伝送容量によつて決まる所定のデータ伝送速度で記憶された伝送画像データＳ１６を順次伝送データD_{TRANS}として引き出して伝送路９に伝送する。

【００１４】これと同時に伝送バツフアメモリ８は残留しているデータ量を検出して当該残留データ量に応じて変化する残量データＳ１７を量子化回路６にフイードバツクして残量データＳ１７に応じて量子化ステツプサイズを制御することにより、伝送画像データＳ１６として発生されるデータ量を調整することにより伝送バツフアメモリ８内に適正な残量（オーバーフロー又はアンダーフローを生じさせないようなデータ量）のデータを維持できるようになされている。

【００１５】因に伝送バツフアメモリ８のデータ残量が許容上限にまで増量して来たとき、残量データＳ１７によつて量子化回路６の量子化ステツプＳＴＰＳ（図１８）のステツプサイズを大きくすることにより、量子化回路６において粗い量子化を実行させることにより伝送画像データＳ１６のデータ量を低下させる。

【００１６】これとは逆に伝送バツフアメモリ８のデータ残量が許容下限値まで減量して来たとき、残量データＳ１７は量子化回路６の量子化ステツプＳＴＰＳのステツプサイズを小さい値になるように制御し、これにより量子化回路６において細かい量子化を実行させるようにすることにより伝送画像データＳ１６のデータ発生量を増大させる。

【００１７】

【発明が解決しようとする課題】ところでかかる構成の動画符号化データ発生装置１から送出される伝送データD_{TRANS}においては、図１９（Ａ）及び（Ｂ）に示すようにフレーム内符号化処理された完全フレーム内処理フレーム（以下これをイントラフレームと呼び、符号「Ａ」で表す）Ａ１、Ａ９、……と、フレーム内符号化処理された前フレーム予測処理フレーム（以下これを予測フレームと呼び、符号「Ｂ」で表す）Ｂ３、Ｂ５、Ｂ７、……及びそれらに応じた補間予測処理フレーム（以下これを補間フレームと呼び、符号「Ｃ」で表す）Ｃ２、Ｃ４、Ｃ６、……が、画像データＶＤの入力フレーム順に伝送されている。

【００１８】ところがこのような伝送データD_{TRANS}を受け、例えば補間フレームＣ２を復号化する際には、図１９（Ｃ）に示すようにイントラフレームＡ１及び予測フレームＢ３が必要になり、動画符号化データの復号化装置としてはイントラフレームＡ１及び予測フレームＢ３を受けるまでの間、補間フレームＣ２を遅延させるメモリやその制御回路が必要になり、その分回路構成が複雑になると共に遅延量が多大になることを避け得なかつた。

【００１９】このため伝送データD_{TRANS}を図１９（Ｃ）に示すように復号化処理に必要な順序で伝送することが考えられ、このような場合例えば伝送データD_{TRANS}はイントラフレームＡ１、Ａ９、……間の８フレーム分（Ａ１、Ｃ２、Ｂ３、Ｃ４、Ｂ５、Ｃ６、Ｂ７、Ｃ８）でなるフレーム群ＧＯＦ１、ＧＯＦ２がＣＤ－ＭＯデイスクの20セクタ分として記録するようになされている。

【００２０】ところがこのようなフレーム順でＣＤ－ＭＯデイスクに記録された動画符号化データについて、例えばフレーム群ＧＯＦ１を編集して書き換えるエデイツト処理を実行し、先頭から順次再生する際には、フレーム群ＧＯＦ１中の第８の補間フレームＣ８は、フレーム群ＧＯＦ１中の新たな第７の予測フレームＢ７とフレーム群ＧＯＦ２中の古い第１のイントラフレームＡ９に基づいて補間されることにより映像信号を得ることができず、結局エデイツト結果を正しく再生できないという問題があつた。

【００２１】本発明は以上の点を考慮してなされたもので、従来の問題を一挙に解決して、編集された動画像符号化データを、簡易な構成で正しく再生することができる符号化データ復号化方法及び装置を提案しようとするものである。

【００２２】

【課題を解決するための手段】かかる課題を解決するため本発明においては、イントラ符号化によつて生成された第１の符号化データと、表示順上で時間的に前にある画像のみを予測画像として用いることが可能な第１のインター符号化によつて生成された第２の符号化データと、表示順上で時間的に前後にある画像を予測画像として用いることが可能な第２のインター符号化によつて生成された第３の符号化データとを含む符号化画像データVD_{REC}が、それぞれいくつかの画像に対応する符号化データを含む複数グループＧＯＦにグループ化され、符号化画像データV_{REC}は、元の先行するグループＧＯＦとは異なるグループＧＯＦに先行されるあるグループＧＯＦを少なくとも１つ含むように編集されてなり、上記あるグループＧＯＦは、元の先行するグループＧＯＦに属する画像を予測画像に用いて第２のインター符号化がなされた画像に対応する符号化データを含み、上記あるグループＧＯＦには編集フラグS_{LPG}が設定されているような符号化画像データVD_{REC}（VD_{PB}）を復号化する符号化データ復号化装置２１において、編集フラグS_{LPG}の設定の有無を検出する手段９０と、編集フラグS_{LPG}が設定されていた場合、上記元の先行するグループＧＯＦに属する画像を上記予測画像に用いて第２のインター符号化された画像を出力しないようにする復号化手段とを設けるようにする。

【００２３】また本発明においては、復号化手段は、元の先行するグループＧＯＦに属する画像を予測画像に用いて第２のインター符号化された画像のかわりに、直前に復号化されたイントラ符号化された画像を出力するようにする。

【００２４】

【発明の実施の形態】以下図面について、本発明を動画符号化データ記録再生装置に適用した場合の実施例を詳述する。

【００２５】（１）動画符号化／復号化装置の全体構成図１及び図２において動画符号化／復号化装置２１は動画符号化装置２１Ａ及び動画復号化装置２１Ｂによつて構成され、動画符号化装置２１Ａは入力映像信号VD_{IN}を入力回路部２２において前処理した後、アナログ／デイジタル変換回路２３において16×16画素分の画素データでなる伝送単位ブロツクデータ、すなわちマクロブロツクＭＢの画素データでなる入力画像データＳ２１を画素データ処理系ＳＹＭ１に送り込むと共に、当該画素データ処理系ＳＹＭ１の各処理段においてマクロブロツクＭＢを単位として画素データが処理されるタイミングにおいて当該処理されるデータに対応する処理情報データがヘツダデータ処理系ＳＹＭ２を介して順次伝送されて行くようになされ、かくして画素データ及びヘツダデー

タがそれぞれ画素データ処理系ＳＹＭ１及びヘツダデータ処理系ＳＹＭ２において並列処理される。

【００２６】この実施例の場合、入力画像データＳ２１として順次送出されて来るマクロブロツクデータは、図３に示すような手法でフレーム画像データＦＲＭから抽出される。第１に入力映像信号VD_{IN}がＱＣＩＦの画サイズ（ 176× 144画素）でなる場合、先ず１枚のフレーム画像データＦＲＭは図３（Ａ１）に示すように１個（水平方向に）×３個（垂直方向に）のブロツクグループＧＯＢに分割され、各ブロツクグループＧＯＢが図３（Ｂ）に示すように11個（水平方向に）×３個（垂直方向に）のマクロブロツクＭＢを含むようになされ、各マクロブロツクＭＢは図３（Ｃ）に示すように16×16画素分の輝度信号データY_{00}～Y_{11}（それぞれ８×８画素分の輝度信号データでなる）及び輝度信号データY_{00}～Y_{11}の全画素データに対応する色信号データでなる色信号データC_b 及びC_r を含んでなる。

【００２７】これに対して第２に入力映像信号VD_{IN}がＣＩＦの画サイズ（252 ×288 画素）でなる場合、１枚のフレーム画像データＦＲＭは図３（Ａ２）に示すように２個（水平方向に）×６個（垂直方向に）のブロツクグループＧＯＢに分割され、各ブロツクグループＧＯＢが図３（Ｂ）に示すように11個（水平方向に）×３個（垂直方向に）のマクロブロツクＭＢを含むようになされ、各マクロブロツクＭＢは図３（Ｃ）に示すように16×16画素分の輝度信号データY_{00}～Y_{11}（それぞれ８×８画素分の輝度信号データでなる）及び輝度信号データY_{00}～Y_{11}の全画素データに対応する色信号データでなる色信号データC_b 及びC_r を含んでなる。

【００２８】かくしてマクロブロツクＭＢごとに送出される入力画像データＳ２１は動き補償回路２５に与えられ、動き補償回路２５はヘツダデータ処理系ＳＹＭ２に対して設けられている動き補償制御ユニツト２６から与えられる動き検出制御信号Ｓ２２に応動して予測前フレームメモリ２７の予測前フレームデータＳ２３と入力画像データＳ２１とを比較して動きベクトルデータMVD(x)及びMVD(y)を検出して動き補償制御ユニツト２６に第１のヘツダデータＨＤ１（図４）のデータとして供給すると共に、動き補償回路本体２５Ａにおいて予測前フレームデータＳ２３に対して動きベクトルデータMVD(x)及びMVD(y)分の動き補償をすることにより予測現フレームデータＳ２４を形成して現在処理しようとしている入力画像データＳ２１でなる現フレームデータＳ２５と共に画像データ符号化回路２８に供給する。

【００２９】ここで動き補償制御ユニツト２６は、図４に示すように、第１のヘツダデータＨＤ１として現在処理しているマクロブロツクごとに、フレーム画像データＦＲＭの伝送順序を表す伝送フレーム番号データTR Counterと、そのブロツクグループＧＯＢ（図３（Ａ１）、（Ａ２））を表すブロツクグループ番号データGOB address と、そのうちのマクロブロツクＭＢを表すマクロブロツク番号データMB addressとを付加することによつて順次画素データ処理系ＳＹＭ１の各処理段に伝送されて行くマクロブロツクＭＢを表示するようになされていると共に、当該処理対象マクロブロツクＭＢの処理ないし処理形式を表すフラグデータFLAGS と、当該マクロブロツクＭＢの動きベクトルデータMVD(x)及びMVD(y)と、その評価値を表す差分データΣ｜Ａ－Ｂ｜とを形成する。

【００３０】フラグデータFLAGS は図５に示すように、最大限１ワード（16ビツト）分のフラグをもち得るようになされ、第０ビツトには、当該処理対象マクロブロツクＭＢについて動き補償モードで処理すべきか否かを表す動き補償制御フラグMC on/off がセツトされる。またフラグデータFLAGS の第１ビツトには、当該処理対象マクロブロツクＭＢをフレーム間符号化モードで処理すべきであるか又はフレーム内符号化モードで処理すべきであるかを表すフレーム間／フレーム内フラグInter/Intra がセツトされる。

【００３１】またフラグデータFLAGS の第２ビツトには、動き補償回路２５のループフイルタ２５Ｂを使用するか否かを表すフイルタフラグFilter on/off が設定される。またフラグデータFLAGS の第３ビツトには、当該処理対象マクロブロツクに含まれるブロツクデータY_{00}～C_r （図３（Ｃ））を伝送すべきであるか否かを表す送信フラグCoded/Not-coded を設定できるようになされている。

【００３２】またフラグデータFLAGS の第４ビツトには、当該処理対象マクロブロツクＭＢを駒落しするか否かを表す駒落しフラグDrop frame flag を設定し得るようになされている。またフラグデータFLAGS の第５ビツトには、当該処理対象マクロブロツクＭＢを強制リフレツシユするか否かを表す強制リフレツシユフラグRefresh on/offを設定できるようになされている。

【００３３】またフラグデータFLAGS の第６ビツトには、マクロブロツクパワー評価フラグMBP appreciateを設定できるようになされている。また差分データΣ｜Ａ－Ｂ｜は、現フレームデータＳ２５の現在処理しようとするマクロブロツクデータＡと、予測前フレームデータＳ２３の検出用動きベクトルによつて補償されたマクロブロツクデータＢとの差分のうちの最小値を表し、これにより検出された動きベクトルの評価をなし得るようになされている。

【００３４】画像データ符号化回路２８はフレーム内符号化モードのとき動き補償回路２５から与えられる現フレームデータＳ２５をそのまま差分データＳ２６として変換符号化回路２９に供給し、これに対してフレーム間符号化モードのとき現フレームデータＳ２５の画素データと予測現フレームデータＳ２４の画素データとの差分でなる差分データＳ２６を変換符号化回路２９に供給する。

【００３５】ヘツダデータ処理系ＳＹＭ２には画像データ符号化回路２８に対応するようにフレーム間／フレーム内符号化制御ユニツト３０が設けられ、動き補償制御ユニツト２６から供給されるヘツダデータＨＤ１及び画像データ符号化回路２８から供給される演算データＳ３１に基づいて、画像データ符号化回路２８の符号化モードを指定するためのフレーム間／フレーム内フラグInter/Intra （図５）及び動き補償回路２５のループフイルタ２５Ｂの動作を制御するためのフイルタフラグFilter on/off （図５）とを得るのに必要なデータを演算して第２のヘツダデータＨＤ２としてフイルタ制御ユニツト３１に送出する。

【００３６】第２のヘツダデータＨＤ２は、図４に示すように、ヘツダデータＨＤ１を構成する伝送フレーム番号データTR Counter～差分データΣ｜Ａ－Ｂ｜をそのまま引き継ぐと共に、フイルタ制御ユニツト３１においてフレーム間／フレーム内符号化モード切換信号Ｓ３３及びフイルタオン／オフ信号Ｓ３４を形成するために必要なパワーデータΣ（Ａ）^{2}（Ｌ）及びΣ（Ａ）^{2}（Ｈ）、Σ（Ａ－Ｂ）^{2}（Ｌ）及びΣ（Ａ－Ｂ）^{2}（Ｈ）、Σ（Ａ－ＦＢ）^{2}（Ｌ）及びΣ（Ａ－ＦＢ）^{2}（Ｈ）、Σ（Ａ）をフレーム間／フレーム内符号化制御ユニツト３０において付加されるようになされている。

【００３７】ここで、パワーデータΣ（Ａ）^{2}（Ｌ）及びΣ（Ａ）^{2}（Ｈ）は現フレームデータＳ２５のマクロブロツク画素データＡの２乗和の下位ビツト及び上位ビツトを表し、パワーデータΣ（Ａ－Ｂ）^{2}（Ｌ）及びΣ（Ａ－Ｂ）^{2}（Ｈ）は現フレームデータＳ２５のマクロブロツク画素データＡとループフイルタ２５Ｂを介さずに形成された予測現フレームデータＳ２４のマクロブロツク画素データＢとの差分Ａ－Ｂの２乗和の下位ビツト及び上位ビツトを表し、パワーデータΣ（Ａ－ＦＢ）^{2}（Ｌ）及びΣ（Ａ－ＦＢ）^{2}（Ｈ）は現フレームデータＳ２５のマクロブロツク画素データＡとループフイルタ２５Ｂを介して形成された予測現フレームデータＳ２４のマクロブロツク画素データＦＢとの差分Ａ－ＦＢの２乗和の下位ビツト及び上位ビツトを表し、パワーデータΣ（Ａ）は現フレームデータＳ２５のマクロブロツク画素データＡの和を表し、それぞれ処理するデータの大きさを評価するためにデータ量をパワー値として表現したもの（２乗和は符号に無関係な値として求めた）である。

【００３８】フイルタ制御ユニツト３１は、フレーム間／フレーム内符号化制御ユニツト３０から渡された第２のヘツダデータＨＤ２と、伝送バツフアメモリ３２から供給される残量データＳ３２とに基づいて、画像データ符号化回路２８に対してフレーム間／フレーム内符号化モード切換信号Ｓ３３を送出すると共に、ループフイルタ２５Ｂに対してフイルタオン／オフ信号Ｓ３４を送出すると共に、当該フイルタオン／オフ信号Ｓ３４の内容を表すフイルタフラグFilter on/off を第２のヘツダデータＨＤ２に付加して第３のヘツダデータＨＤ３としてスレシヨルド制御ユニツト３５に渡す。ここでフイルタ制御ユニツト３１は第１に、フレーム間符号化処理をした場合の伝送データ量の方がフレーム内符号化処理をした場合の伝送データ量より大きくなつたとき画像データ符号化回路２８をフレーム内符号化モードに制御する。

【００３９】またフイルタ制御ユニツト３１は第２に、フレーム間符号化モードで処理をしている状態においてループフイルタ２５Ｂにおける処理を受けた予測現フレームデータＳ２４より当該処理を受けない予測現フレームデータＳ２４の方が差分値が小さい場合には、フイルタオン／オフ信号Ｓ３４によつてフイルタリング動作をさせないようにループフイルタ２５Ｂを制御する。

【００４０】またフイルタ制御ユニツト３１は第３に、強制リフレツシユモードになつたとき、フレーム間／フレーム内符号化モード切換信号Ｓ３３によつて画像データ符号化回路２８をフレーム内符号化モードに切り換える。さらにフイルタ制御ユニツト３１は第４に、伝送バツフアメモリ３２から供給される残量データＳ３２に基づいて伝送バツフアメモリ３２がオーバーフローするおそれがある状態になつたとき、これを検出して駒落し処理をすべきことを命令するフラグを含んでなる第３のヘツダデータＨＤ３をスレシヨルド制御ユニツト３５に送出する。

【００４１】かくして画像データ符号化回路２８は現フレームデータＳ２５と予測現フレームデータＳ２４との差分が最も小さくなるようなモードで符号化してなる差分データＳ２６を変換符号化回路２９に供給する。

【００４２】第３のヘツダデータＨＤ３は、図４に示すように、ヘツダデータＨＤ２から伝送フレーム番号データTR Counter～動きベクトルデータMVD(x)及びMVD(y)を引き継ぐと共に、フイルタ制御ユニツト３１においてブロツクデータY_{00}～C_r に対応する６ビツト分のフイルタフラグFilter on/off を付加される。

【００４３】変換符号化回路２９はデイスクリートコサイン変換回路でなりデイスクリートコサイン変換後の係数値を６個のブロツクY_{00}、Y_{01}、Y_{10}、Y_{11}、C_b、C_rごとにジグザグスキヤンしてなる変換符号化データＳ３５として伝送ブロツク設定回路３４に送出する。

【００４４】伝送ブロツク設定回路３４は変換符号化データＳ３５として送出されて来る６個のブロツクデータY_{00}～C_r （図３（Ｃ））について、それぞれ先頭の係数データからｎ個までの２乗和を演算して当該演算結果をパワー検出データＳ３６としてスレシヨルド制御ユニツト３５に渡す。

【００４５】このときスレシヨルド制御ユニツト３５は各ブロツクデータY_{00}～C_r ごとにパワー検出データＳ３６を所定のスレシヨルドと比較し、パワー検出データＳ３６が当該スレシヨルドより小さいとき当該ブロツクデータの伝送を許容せず、これに対して大きいとき許容

することを表す６ビット分の伝送可否データＣＢＰＮを形成してこれをフイルタ制御ユニット３１から渡された第３のヘッダデータＨＤ３に付加して第４のヘッダデータＨＤ４として量子化制御ユニット３６に渡すと共に、伝送ブロック設定回路３４から対応するブロックデータY_{00}～C_rを量子化回路３７に送信ブロックパターン化データＳ３７として送出させる。

【００４６】ここで第４のヘッダデータＨＤ４は図４に示すように、ヘッダデータＨＤ３の伝送フレーム番号データTR Counter～フイルタフラグFilter on/off をそのまま引き継ぐと共に、スレショルド制御ユニット３５においてブロックY_{00}～C_rに対応して発生する６ビット分の送信可否フラグＣＢＰＮが付加される。

【００４７】量子化制御ユニット３６はスレショルド制御ユニット３５から渡された第４のヘッダデータＨＤ４と、伝送バッファメモリ３２から送出される残量データＳ３２とに基づいて、量子化ステップサイズ決定処理を実行して得られる量子化ステップサイズ制御信号Ｓ３８を量子化回路３７に与え、これにより量子化回路３７をマクロブロックＭＢに含まれるデータに適応した量子化ステップサイズで量子化処理させ、その結果量子化回路３７の出力端に得られる量子化画像データＳ３９を可変長符号化回路３８に供給させる。

【００４８】これと共に量子化制御ユニット３６は、図４に示すように、第５のヘッダデータＨＤ５として、ヘッダデータＨＤ４に基づいてブロックデータ　Y_{00}～C_r　（図３（Ｃ））にそれぞれ対応するフラグデータFLAGS 及び動きベクトルデータMVD(x)及びMVD(y)に分離してこれを直列に配列させたデータを形成して可変長符号化回路３８及び逆量子化回路４０に渡す。

【００４９】ここで、ヘッダデータＨＤ５は、図４に示すように、ヘッダデータＨＤ４のうち伝送フレーム番号データTR Counter～マクロブロック番号データMB addressをそのまま引き継ぐと共に、量子化制御ユニット３６において量子化サイズデータQNT と、ブロックデータY_{00}～C_rに対するフラグデータFLAGS 、動きベクトルデータMVD(x)及びMVD(y)を付加する。

【００５０】可変長符号化回路３８はヘッダデータＨＤ５及び量子化画像データＳ３９を可変長符号化処理して伝送画像データＳ４０を形成し、これを伝送バッファメモリ３２に供給する。可変長符号化回路３８はブロックデータY_{00}～C_rを可変長符号化する際に、対応するフラグデータFLAGS に基づいて「駒落し」、又は「送信不可」が指定されているとき、当該ブロックデータを伝送画像データＳ４０として送出させずに捨てるような処理をする。

【００５１】伝送バッファメモリ３２は伝送画像データＳ４０を溜め込んで行くと共に、これを所定の伝送速度で読み出してマルチプレクサ４１において音声データ発生装置４２から送出される伝送音声データＳ４１と合成して動画符号化データVD_{REC}としてＣＤ－ＭＯ装置に送出する。

【００５２】逆量子化回路４０は量子化回路３７から送出される量子化画像データＳ３９をヘッダデータＨＤ５に基づいて逆量子化した後、当該逆量子化データＳ４２を逆変換符号化回路４３に供給することにより逆変換符号化データＳ４３に変換させた後デコーダ回路４４に供給させ、かくして伝送画像データＳ４０として送出された画像情報を表す符号化差分データＳ４４を予測前フレームメモリ２７に供給させる。このとき予測前フレームメモリ２７は、符号化差分データＳ４４を用いてそれまで保存していた予測前フレームデータを修正演算して新たな予測前フレームデータとして保存する。

【００５３】かくして図１の構成の動画符号化装置２１Ａによれば、ヘッダデータ処理系ＳＹＭ２から供給されるヘッダ情報に基づいて画素データ処理系ＳＹＭ１において画素データがマクロブロック単位でパイプライン処理されて行くのに対して、これと同期するようにヘッダデータ処理系ＳＹＭ２においてヘッダデータを受け渡して行くようにすることにより、ヘッダデータ処理系ＳＹＭ２の各処理段において必要に応じてヘッダデータを付加又は削除することにより画素データを必要に応じて適応処理できる。

【００５４】動画復号化装置２１Ｂは図２に示すように、ＣＤ－ＭＯ装置から再生される動画符号化データVD_{PB}をデマルチプレクサ５１を介して伝送バッファメモリ５２に受けると共に、伝送音声データＳ５１を音声データ受信装置５３に受ける。

【００５５】伝送バッファメモリ５２に受けた画像データは可変長逆変換回路５４において受信画像データＳ５２及びヘッダデータＨＤ１１に分離され、逆量子化回路５５において逆量子化データＳ５３に逆量子化された後、逆変換符号化回路５６においてデイスクリート逆変換処理されて逆変換符号化データＳ５４に逆変換される。

【００５６】この逆変換符号化データＳ５４は逆量子化回路５５において形成されたヘッダデータＨＤ１２と共にデコーダ回路５７に与えられ、符号化差分データＳ５５としてフレームメモリ５８に蓄積される。

【００５７】かくしてフレームメモリ５８には符号化差分データＳ５５に基づいて伝送されてきた画像データが復号化され、当該復号化画像データＳ５６がデイジタル／アナログ変換回路５９においてアナログ信号に変換された後、出力回路部６０を介して出力映像信号VD_{OUT}として送出される。

【００５８】（２）実施例による動画符号化データの記録順序

図１及び図２との対応部分に同一符号を付して示す図６において、７０は全体として本発明による符号化データ復号化方法を適用した動画符号化データ記録再生装置を

示す。この動画符号化データ記録再生装置７０の場合、入力映像信号VD_{IN}が上述した動画符号化／復号化装置２１を通じて高能率符号化され、この結果得られる動画符号化データVD_{REC}がＣＤ－ＭＯ装置７１に入力されてＣＤ－ＭＯデイスク（図示せず）に記録される。

【００５９】逆にＣＤ－ＭＯ装置７１から得られる再生信号が動画符号化データVD_{PB}として、動画符号化／復号化装置２１に入力され、この結果動画符号化データVD_{PB}を復号化して得られる出力映像信号VD_{OUT}が送出される。この動画符号化データ記録再生装置７０の場合、動画符号化／復号化装置２１はバス７２を通じてＣＰＵ（中央処理ユニツト）を含む記録再生制御回路７３に接続され、この記録再生制御回路７３によつて入力映像信号VD_{IN}の符号化及び動画符号化データVD_{PB}の復号化が制御される。

【００６０】これに加えてＣＤ－ＭＯ装置７１はＳＣＳＩ(small computer system interface)を内蔵し、ＳＣＳＩバス７４、ＳＣＳＩ制御回路７５及びバス７２を通じて、記録再生制御回路７３に接続され、これにより記録再生制御回路７３によつて記録再生動作が制御される。

【００６１】ここで動画符号化／復号化装置２１から送出される動画符号化データVD_{REC}は、図7に示すような階層（レイヤ）構造を有するフオーマツトでＣＤ－ＭＯ装置７１に入力され、またＣＤ－ＭＯ装置７１から同様のフオーマツトで動画符号化データVD_{PB}として動画符号化／復号化装置２１に入力される。

【００６２】すなわち動画符号化データVD_{REC}及びVD_{PB}においては、フレーム群レイヤとして入力画像信号VD_{IN}の８フレーム分に対応したデータを１フレーム群ＧＯＦとして、当該１フレーム群の開始を表すフレーム群スタートコード(GOFSC)、直前のＧＯＦとの連続関係を表すリンクフラグ(LPG)、伝送するフレームの水平、垂直サイズや水平及び垂直方向の画素数比等を表すデータ(HORSIZE、VERSIZE、HVPRAT)、伝送フレームのレートを表すデータ(RATE)、１フレーム分の画像データでなるピクチヤレイヤのデータ（P.data）の８フレーム分及びスタツフイングビツト(TSB)から構成されている（図７（Ａ））。

【００６３】ピクチヤレイヤのデータ（P.data）の１フレーム分は、１フレームの開始を表すフレームスタートコード(PSC)、フレーム番号(TR)、拡張情報を表すデータ(PEI、PSPARE)及びブロツク単位の画像データでなるブロツクグループレイヤのデータ(GOB data)の１フレーム分から構成されている（図７（Ｂ））。

【００６４】ブロツクグループレイヤのデータ(GOB data)の１ブロツクグループ分は、１ブロツクグループの開始を表すブロツクグループスタートコード(GBSC)、ブロツクグループのアドレスデータ（GN）、ブロツクグループ単位の再量子化ステツプサイズに関するデータ(GQUANT)、拡張情報を表すデータ(GEI、GSPARE)及びマクロブロツクレイヤのデータ(MB data)の１ブロツクグループ分から構成されている（図７（Ｃ））。

【００６５】マクロブロツクレイヤのデータ(MB data)の１マクロブロツク分は、マクロブロツクのアドレスを表すデータ(MBA)、マクロブロツクのタイプを表すデータ(MTYPE)、マクロブロツクにおける再量子化ステツプサイズのデータ(MQUANT)、マクロブロツク毎の動きベクトルのデータ(MVD1、MVD2)、マクロブロツク内のブロツクパターンのデータ(CBP)及びブロツクレイヤのデータ(Block data)の１マクロブロツク分から構成されている（図７（Ｄ））。

【００６６】ブロツクレイヤのデータ(Block data)の１ブロツク分は、所定数の係数データ(TCOEF)とブロツクレイヤの終わりを表すデータ(EOB)から構成されている（図７（Ｅ））。

【００６７】ここでこの実施例による動画符号化装置２１Ａにおいては、図８（Ａ）に示す従来同様の入力画像信号VD_{IN}に基づくフレーム記録順序に代え、図８（Ｂ）に示すように動画復号化装置２１Ｂ側の復号化処理に応じたフレーム記録順序でなる動画符号化データVD_{REC}をＣＤ－ＭＯ装置７１に送出し記録するようになされ、ＣＤ－ＭＯ装置７１から再生された動画符号化データVD_{PB}がこのフレーム記録順序で動画復号化装置２１Ｂに入力される。

【００６８】このようにすれば、例えば補間フレームＣ２を復号化する際、復号化に必要なイントラフレームＡ１及び予測フレームＢ３がすでに入力され、また例えば補間フレームＣ４を復号化する際、復号化に必要な予測フレームＢ３及びＢ５がすでに入力され、これにより動画復号化装置２１Ｂは直ちに補間フレームＣ２又はＣ４の復号化処理を実行することができる。

【００６９】ここでこの実施例の動画符号化装置２１Ａの場合、図９に示すように、動き補償回路２５に内蔵されたフレーム順並べ替え回路８０を用いて、入力画像信号VD_{IN}に基づくフレーム記録順序を上述のように復号化処理に応じたフレーム記録順序に並べ替えるようになされている。

【００７０】このフレーム順並べ替え回路８０においては、第１～第３の１フレーム遅延回路８１、８２及び８３を有して構成されており、フレーム順の並べ替え処理に加えて、動きベクトルの検出処理を実行し得るようになされている。すなわち入力画像信号VD_{IN}は入力回路部２２及びアナログデイジタル変換回路２３において所定の処理が施され、この結果得られる入力画像データＳ２１が第１の１フレーム遅延回路８１に入力される。

【００７１】この第１の１フレーム遅延回路８１から送出された第１の遅延データS_{D1}は、第２の１フレーム遅延回路８２に入力されると共に、第１のフレーム選択回路８４の第１の入力端ａに入力される。また第２の１フ

レーム遅延回路８２から送出される第２の遅延データS_{D2}は第３の１フレーム遅延回路８３を通じ、さらに１フレーム分遅延され第３の遅延データS_{D3}として第１のフレーム選択回路８４の第２の入力端ｂに入力される。

【００７２】これによりこのフレーム順並べ替え回路８０においては、１フレーム毎のタイミングで順次入力画像データＳ２１を入力すると共に、これに同期したフレームパルスＦＰのタイミングで順次第１のフレーム選択回路８４の第１又は第２の入力端ａ又はｂを選択制御することにより、フレーム順の並べ替え処理を実行し得るようになれている。

【００７３】なお入力画像データＳ２１及び第１の遅延データS_{D1}は第１の動きベクトル検出回路８５に入力され、この結果得られる入力画像データＳ２１及び第１の遅延データS_{D1}間の動きベクトルデータD_{MV1}及び差分データD_{DR1}が動き補償ユニツト２６に送出される。

【００７４】また第２及び第３の遅延データS_{D2}及びS_{D3}が第２のフレーム選択回路８６の第１及び第２の入力端ａ及びｂに入力され、フレームパルスＦＰのタイミングで何れか一方が選択されて第２の動きベクトル検出回路８７に入力される。この第２の動きベクトル検出回路８７には、これに加えて第１の遅延データS_{D1}が入力されており、この結果得られる第１の遅延データS_{D1}及び第２又は第３の遅延データS_{D2}又はS_{D3}間の動きベクトルデータD_{MV2}及び差分データD_{DR2}が動き補償ユニツト２６に送出される。かくして入力画像信号VD_{IN}について高能率符号処理して動画符号化データVD_{REC}を得る際に、入力画像信号VD_{IN}に基づくフレーム順序に代えて、復号化処理に応じたフレーム順序に並べ替えるようにしたことにより、復号化処理側の回路構成及び制御を簡略かつ効率化し得る動画符号化データ伝送方法を実現できる。

【００７５】（３）実施例による動画符号化データのエデイツト処理

ここでこの動画符号化データ記録再生装置７０の記録再生制御回路７３は、外部から入力されたエデイツト命令に応動して、図１０に示すエデイツト処理手順ＲＴ０を実行し、これにより図１１に示すように、ＣＤ－ＭＯデイスク上の20セクタ毎に１フレーム群分記録された動画符号化データVD_{REC}を１フレーム群単位で書換え、かくしてエデイツト処理を実行するようになされている。

【００７６】なおこの実施例の場合動画符号化データVD_{IN}においては、図８（Ｃ）に示すようにイントラフレームから続く補間フレーム及び予測フレームＡ１、Ｃ２、Ｂ３、Ｃ４、Ｂ５、Ｃ６、Ｂ７、Ｃ８の８フレーム分を１フレーム群ＧＯＦ１、ＧＯＦ２、……とした従来の方法に代え、イントラフレームＡ１の直前の補間フレームＣ０に続くイントラフレーム、補間フレーム及び予測フレームＣ０、Ａ１、Ｃ２、Ｂ３、Ｃ４、Ｂ５、Ｃ６、Ｂ７を１フレーム群ＧＯＦ１１、ＧＯＦ１２、……として伝送するようになされている。

【００７７】このようにして１フレーム群ＧＯＦ１１、ＧＯＦ１２、……内のフレームが、他のフレーム群ＧＯＦ１１、ＧＯＦ１２、……に含まれてしまう不都合を未然に防止し得るようになされている。実際上記録再生制御回路７３は、図１０に示すエデイツト処理手順ＲＴ０から入つて次のステツプＳＰ１においてエデイツト命令の解析処理を実行する。

【００７８】ここで例えば図１１に示す第３及び第４のフレーム群ＧＯＦ３及びＧＯＦ４の40セクタ分について、新たな動画符号化データVD_{REC}と書き換えることが指示されると、記録再生制御回路７３は次のステツプＳＰ２に移る。

【００７９】このステツプＳＰ２において、記録再生制御回路７３はＳＣＳＩ制御回路７５にエデイツト命令に応じた制御命令を送出し、これによりＳＣＳＩバス７４を通じてＣＤ－ＭＯデイスクの書き換え制御を実行する。なおこのときバス７２を通じた記録再生制御回路７３の制御によつて、新たな２フレーム群ＧＯＦ３Ｎ、ＧＯＦ４Ｎ分の入力映像信号VD_{IN}が動画符号化／復号化装置２１を通じて動画符号化データVD_{REC}としてＣＤ－ＭＯ装置７１に入力される。

【００８０】続いて記録再生制御回路７３は次のステツプＳＰ３を実行し、書換え処理が終了したか否かを判断し、ここで否定結果を得るとステツプＳＰ２に戻りＣＤ－ＭＯデイスクの書き換え制御を継続し、やがて肯定結果を得るとステツプＳＰ４に移る。このステツプＳＰ４において記録再生制御回路７３は、エデイツトしたＣＤ－ＭＯデイスクの先頭セクタ（この実施例の場合、第40セクタでなる）の内容を読む。

【００８１】これにより記録再生制御回路７３は次のステツプＳＰ５において、先頭セクタの先頭から25ビツト目に存在するフレーム群レイヤのリンクフラグ(LPG)をエデイツトフラグとして設定し、これをＣＤ－ＭＯデイスクの読み出したセクタ位置に書き込む。

【００８２】続いて記録再生制御回路７３はステツプＳＰ６において、エデイツトしたＣＤ－ＭＯデイスクの最終セクタに続くセクタ（この実施例の場合、第80セクタでなる）の内容を読む。これにより記録再生制御回路７３は次のステツプＳＰ７において、当該セクタの先頭から25ビツト目に存在するフレーム群レイヤのリンクフラグ(LPG)を上述と同様にエデイツトフラグとして設定し、これをＣＤ－ＭＯデイスクの読み出したセクタ位置に書き込み、次のステツプＳＰ８において当該エデイツト処理手順ＲＴ０を終了する。

【００８３】実際上このようにしてエデイツトされてＣＤ－ＭＯデイスク上に記録された動画符号化データVD_{REC}は、記録再生制御回路７３の制御によつて読み出され、この結果再生信号として得られる動画符号化データVD_{PB}が動画符号化／復号化装置２１に入力される。

【００８４】（４）実施例によるエデイツトされた符号化データの復号化方法
この実施例の場合動画復号化装置２１Ｂにおいては、図１２に示すようにデコーダ回路５７に含まれるフレーム順逆並べ替え回路９０を用いて、復号化処理に応じたフレーム順序から入力画像信号VD_{IN}に基づくフレーム順序に並べ替える逆並べ替え処理を実行すると共に、フレーム群レイヤのリンクフラグ(LPG) に設定されたエデイツトフラグを参照してエデイツト再生処理を実行するようになされている。
【００８５】すなわち逆変換符号化回路５６から送出される逆変換符号化データＳ５４は、デコーダ回路５７のセレクタ回路９１の第１の入力端Ａに直接入力されると共に、フレームメモリ９２を通じて例えば２フレーム分遅延されて第２に入力端Ｂに入力される。
【００８６】このセレクタ回路９１及びフレームメモリ９２はそれぞれフレームパルスＦＰのタイミングで動作し、これにより復号化処理に応じたフレーム順序（図８（Ｂ））から入力画像信号VD_{IN}に基づくフレーム順序（図８（Ａ））への逆並べ替え処理を実行するようになされている。
【００８７】なおこのセレクタ回路９１には、逆量子化回路５５において形成されたヘツダデータＨＤ１２のうちリンクフラグ(LPG) に設定されたエデイツトフラグに応じたエデイツトフラグ信号S_{LPG}が入力され、このエデイツトフラグ信号S_{LPG}が設定されていることを表すときのみセレクト動作を中断し、次に到来するフレームをそのまま出力する。
【００８８】このようにして例えばエデイツトされた新たなフレーム群ＧＯＦＮ３、ＧＯＦＮ４中のフレームデータに対して、古いフレーム群ＧＯＦ１、ＧＯＦ２、ＧＯＦ５、……中のフレームデータが混入して再生画像に乱れが生じるおそれを未然に防止し得るようになれている。因に図８（Ｄ）に示すフレーム群ＧＯＦ１２にエデイツトフラグが設定されている場合には、逆並べ替え後の第８の補間フレームＣ８に代えて、イントラフレームＡ９が２フレーム分出力される。
【００８９】以上の構成によれば、エデイツトされたフレーム群及びエデイツト直後のフレーム群のリンクフラグ(LPG) にエデイツトフラグS_{LPG}が設定されてなる編集後の動画符号化データを復号化する際に、エデイツトフラグS_{LPG}の有無を検出し、エデイツトフラグS_{LPG}が設定されている場合には元の先行するフレーム群に属する画像を予測画像に用いて補間フレームを出力せずに、直前に復号化したイントラフレームを出力するようにして復号化するようにしたことにより、編集後の動画符号化データを正しく再生することができる。
【００９０】（５）他の実施例
（５－１）上述の実施例においては、記録された動画符号化データについてフレーム群単位でエデイツト処理を実行する際、エデイツトされたフレーム群及びエデイツト直後のフレーム群のリンクフラグ(LPG) にエデイツトフラグを設定した場合について述べたが、これに代え、例えばエデイツト処理したフレーム群を記録再生制御回路７３が記憶している場合には、エデイツト直後のフレーム群のリンクフラグ(LPG) にエデイツトフラグを設定するのみで良好にエデイツト処理を実行し得る。
【００９１】（５－２）また上述の実施例においては、記録された動画符号化データについてフレーム群単位でエデイツト処理を実行する際、エデイツトされたフレーム群及びエデイツト直後のフレーム群のリンクフラグ(LPG) にエデイツトフラグを設定した場合について述べたが、これに代え、動画符号化装置２１Ａ側でピクチヤレイヤのフレーム番号(TR)に、記録再生制御回路７３で発生した所定の乱数から始まる連番を順次付加し、エデイツト再生処理時にこのフレーム番号(TR)の不連続を検出したタイミングで上述したエデイツトフラグ信号S_{LPG}と同様の信号を発生するようにすれば、上述の実施例と同様の効果を実現できる。
【００９２】因に、この場合フレーム番号(TR)の不連続は、図１３に示すような不連続検出回路９５で検出される。すなわちこの不連続検出回路９５においては、上述のフレーム順逆並べ替え回路９０に併設され、逆変換符号化データＳ５４に含まれるピクチヤレイヤのフレーム番号(TR)が比較回路９６及びラツチ回路９７に入力される。
【００９３】このラツチ回路９７はフレームパルスＦＰのタイミングでラツチ動作を実行し、この結果１フレーム分遅延したフレーム番号(TR)が加算回路９８に入力されて値「１」が加算され、比較フレーム番号C_{TR}として比較回路９６に入力される。これにより比較回路９６は、フレーム番号(TR)及び比較フレーム番号C_{TR}の値を比較し、両者が不一致のとき論理「Ｈ」レベルでなる不一致検出信号を発生し、これをアンド回路９９に送出する。
【００９４】アンド回路９９にはこれに加えて、先頭フレーム群の再生処理のとき論理「Ｈ」レベルを有する先頭フレーム群信号ＧＯＦＦが反転回路１００を通じて反転して入力されており、これにより先頭フレーム群の再生処理のとき、不一致検出信号を論理「Ｌ」レベルに制御し、それ以外のとき不一致検出信号に応じた論理レベルを有するエデイツトフラグ信号S_{LPG}をフレーム順逆並べ替え回路９０のセレクタ回路９１に送出する。
【００９５】（５－３）さらに上述の実施例においては、動画符号化データについてフレーム群単位でエデイツト処理を実行する際、エデイツトフラグやフレーム番号(TR)を用いてエデイツトされたフレーム群を識別するようにしたが、フレーム群の各補間フレームＣについてフレーム群内のイントラフレームＡ及び予測フレームＢのみを用いて補間し、フレーム群内で各フレームの関係

が完結するようにすれば、画質について多少の劣化が生じても容易かつ自在にエデイツト処理を実行し得る動画符号化データ記録方法を実現できる。

【００９６】（５－４）さらに上述の実施例においては、イントラフレームＡ、予測フレームＢ及び補間フレームＣが、図８（Ａ）に示すように配置された動画符号化データのフレーム順序を並べ替えるようにした場合について述べたが、動画符号化データのフレーム配置はこれに限らず、図１４（Ａ）や図１５（Ａ）に示すような場合でも、要は図１４（Ｂ）や図１５（Ｂ）に示すように復号化側の処理順に応じたフレーム順に並べ替えて伝送するようにすれば、上述の実施例と同様の効果を実現できる。

【００９７】またこの場合図１４（Ｃ）や図１５（Ｃ）に示すフレーム群ＧＯＦ１、ＧＯＦ２、……の配置に代え、図１４（Ｄ）や図１５（Ｄ）に示すようなフレーム群ＧＯＦ２１、ＧＯＦ２２、……やＧＯＦ３１、ＧＯＦ３２、……の配置にすれば、エデイツト処理についても上述の実施例と同様の効果を実現できる。

【００９８】（５－５）さらに上述の実施例においては、映像信号を高能率符号化してＣＤ－ＭＯデイスクに記録し再生する場合について述べたが、記録媒体はこれに限らず、他の光デイスクや磁気デイスク、磁気テープ等に広く適用して好適なものである。

【００９９】（５－６）さらに上述の実施例においては、映像信号を高能率符号化してＣＤ－ＭＯデイスクに記録し再生する場合について述べたが、本発明はこれに限らず、映像信号を高能率符号化して伝送する動画符号化データ伝送方法に広く適用して好適なものである。

【０１００】

【発明の効果】上述のように本発明によれば、編集フラグの有無を検出し、編集フラグが設定されていた場合には先行するグループに属する画像を予測画像に用いて、表示順上で時間的に前後にある画像を予測画像として用いることが可能なインター符号化によつて生成された画像を出力しないようにしたことにより、編集後の符号化画像データを正しく復元し得る符号化データ復号化方法及び装置を実現できる。

【図面の簡単な説明】

【図１】実施例による動画符号化装置の構成を示すブロツク図である。

【図２】実施例による動画復号化装置の構成を示すブロツク図である。

【図３】フレーム画像データの構成を示す略線図である。

【図４】ヘツダデータ処理系の詳細構成を示すブロツク図である。

【図５】フラグデータの構成を示す略線図である。

【図６】実施例による動画符号化データ記録再生装置の構成を示すブロツク図である。

【図７】記録再生データのフオーマツトの説明に供する略線図である。

【図８】実施例による動画符号化データの記録順序の説明に供する略線図である。

【図９】フレーム順並べ替え回路の構成を示すブロツク図である。

【図１０】エデイツト処理の説明に供するフローチヤートである。

【図１１】ＣＤ－ＭＯデイスクの記録領域の説明に供する略線図である。

【図１２】フレーム順逆並べ替え回路の構成を示すブロツク図である。

【図１３】他の実施例による不連続検出回路を示すブロツク図である。

【図１４】他の実施例による動画符号化データの記録順序の説明に供する略線図である。

【図１５】他の実施例による動画符号化データの記録順序の説明に供する略線図である。

【図１６】フレーム内／フレーム間符号化処理の説明に供する略線図である。

【図１７】従来の動画符号化データ発生装置を示すブロツク図である。

【図１８】その量子化ステツプを示す特性曲線図である。

【図１９】従来の動画符号化データの記録順序の説明に供する略線図である。

【符号の説明】

２１……動画符号化／復号化装置、２１Ａ……動画符号化装置、２１Ｂ……動画復号化装置、２５……動き補償回路、２６……動き補償制御ユニツト、２７……予測前フレームメモリ、２８……画像データ符号化回路、２９……変換符号化回路、３０……フレーム間／フレーム内符号化制御ユニツト、３１……フイルタ制御ユニツト、３２……伝送バツフアメモリ、３４……伝送ブロツク設定回路、３５……スレシヨルド制御ユニツト、３６……量子化制御ユニツト、３７……量子化回路、３８……可変長符号可回路。

【図1】

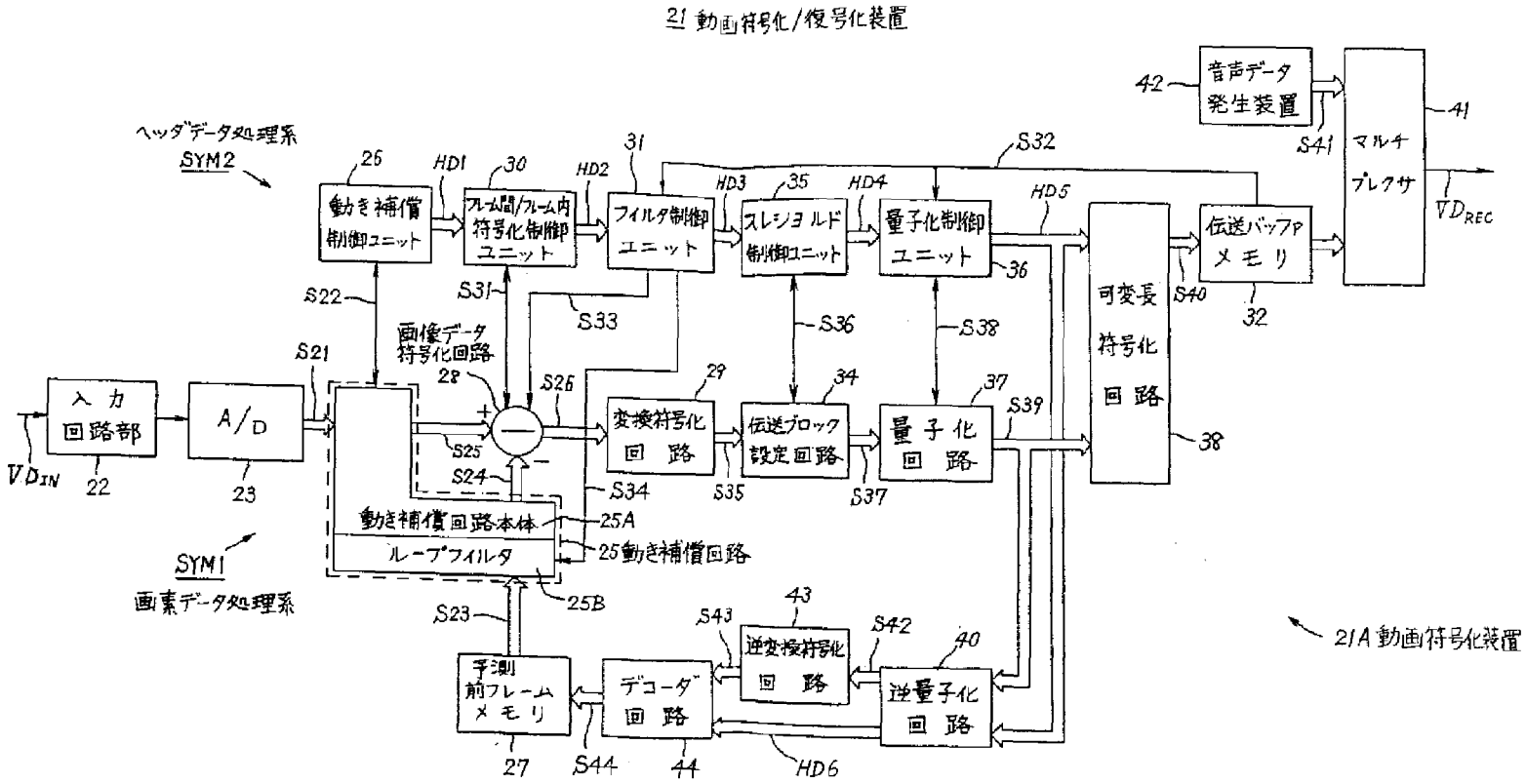


図1　動画符号化装置の構成

【図2】

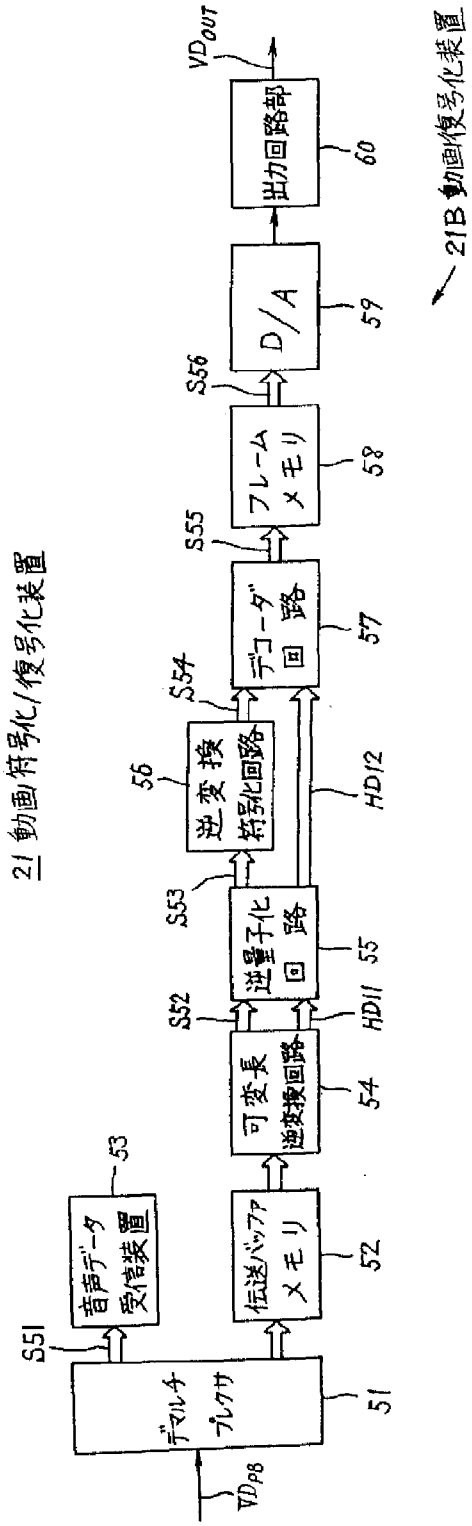


図2　動画復号装置の構成

【図3】

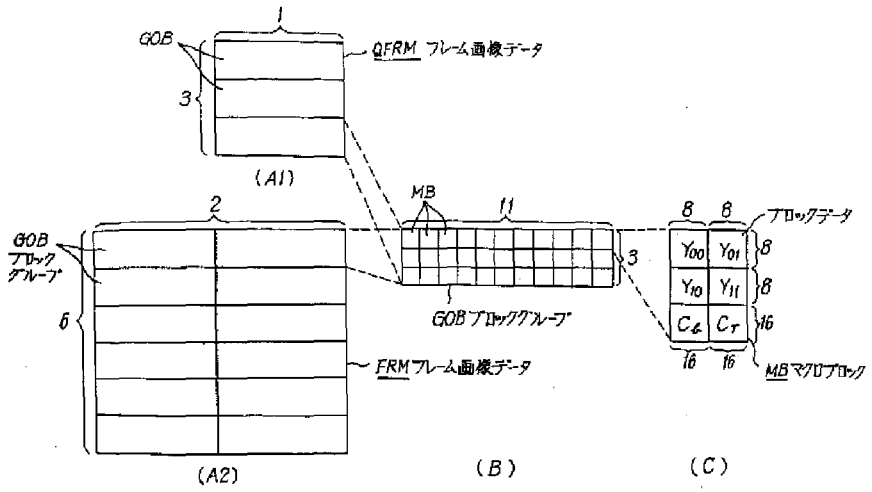


図3　マクロブロックデータの構成

【図6】

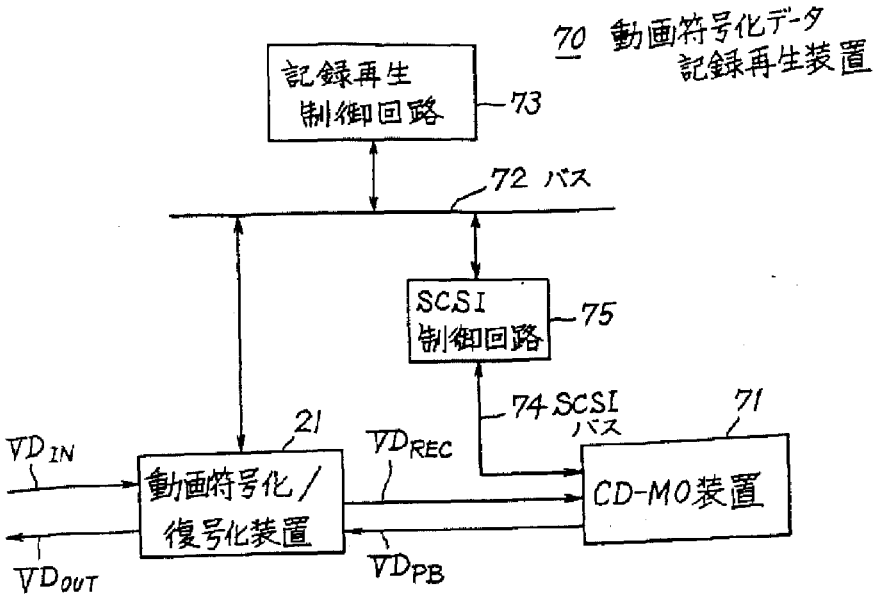


図6　実施例の動画符号化データ記録再生装置の全体構成

【図4】

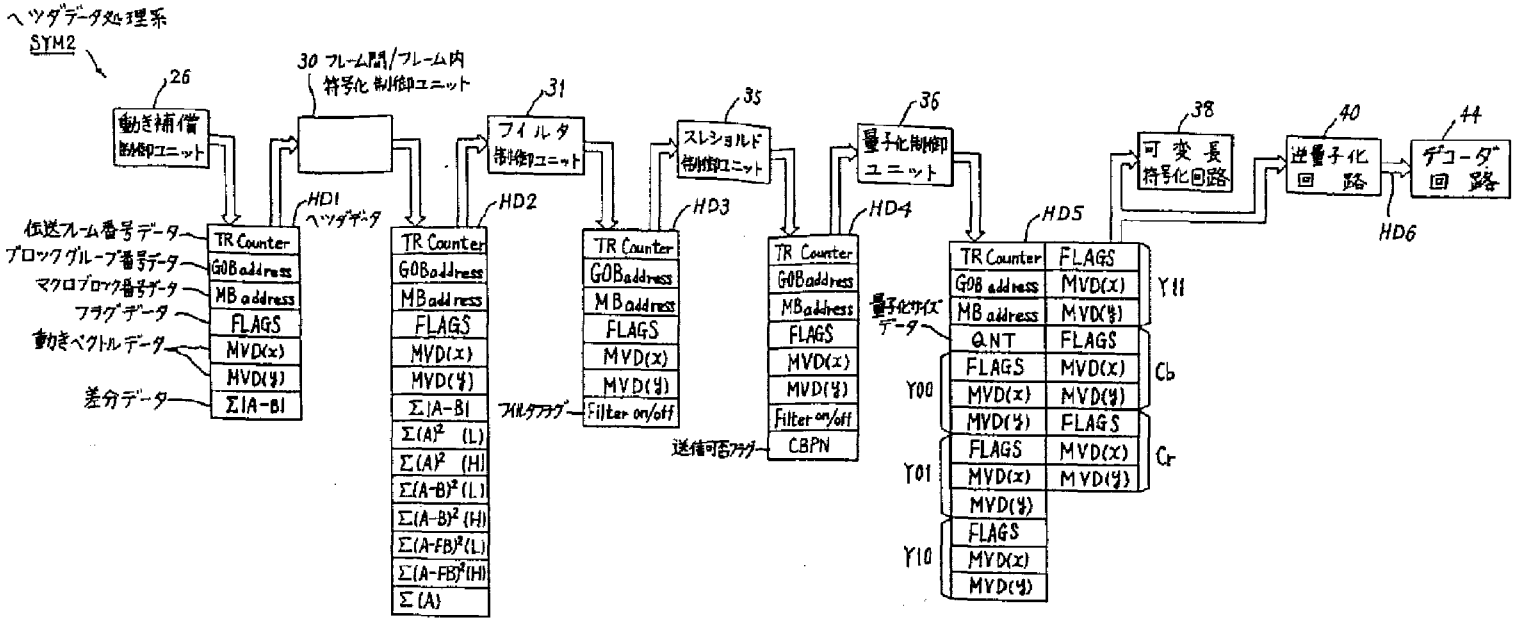


図4　ヘツダデータ処理系の詳細

【図5】

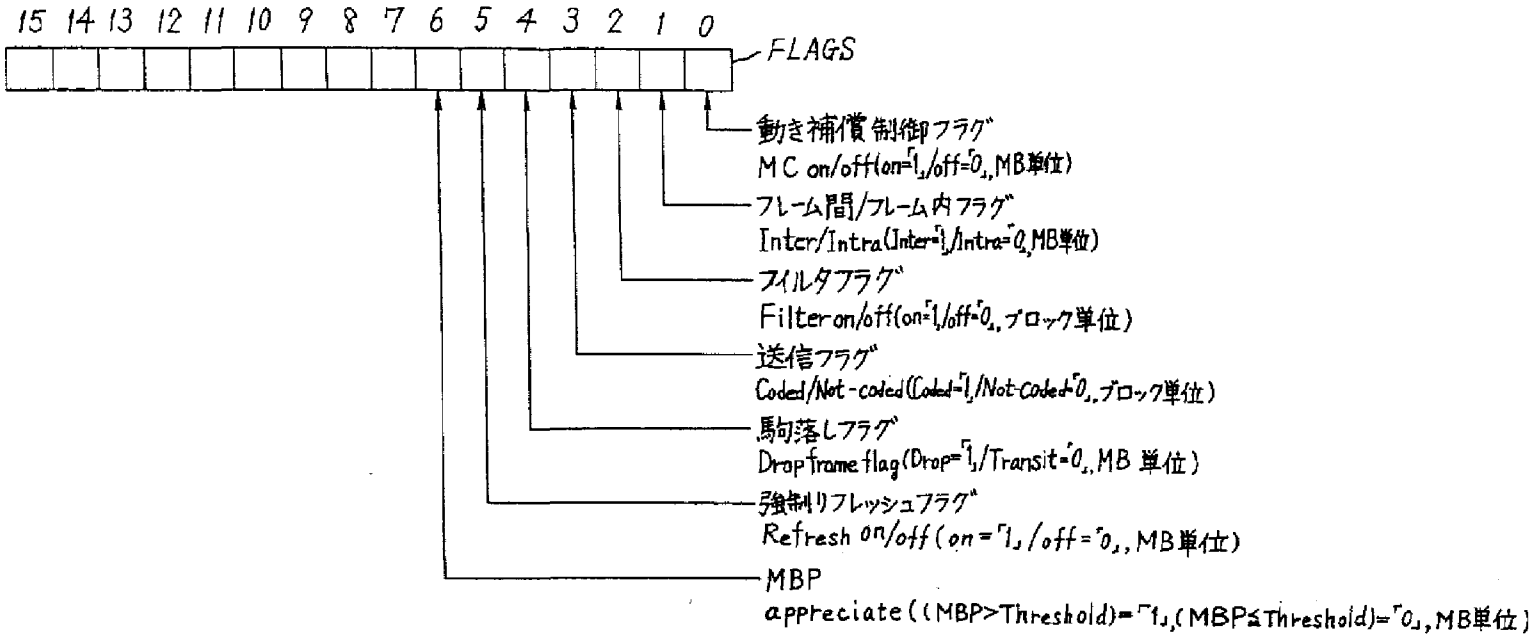


図5　フラグデータの構成

【図11】

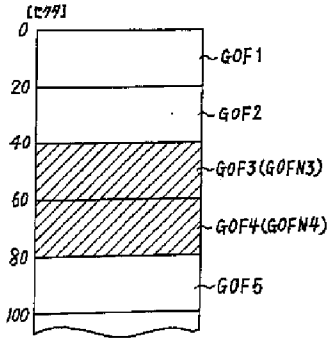


図11　CD－MOディスクの記録領域

【図8】

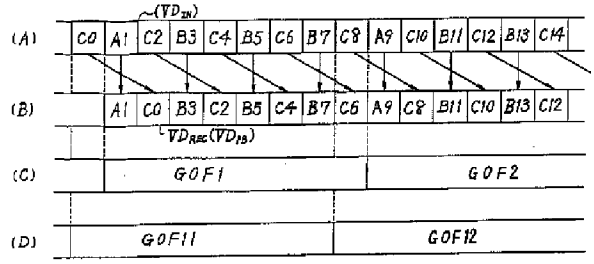


図8　実施例の記録順序

【図7】

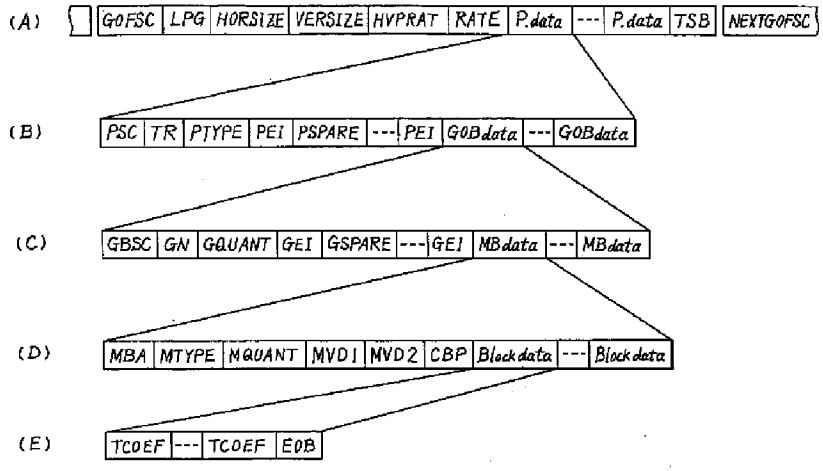


図7　記録再生データのフォーマット

【図12】

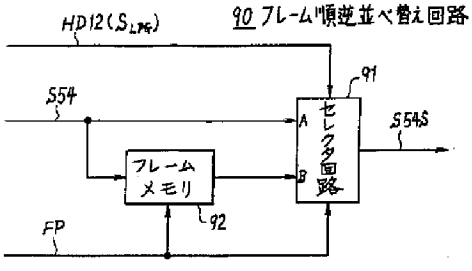


図12　フレーム順逆並べ替え回路

【図13】

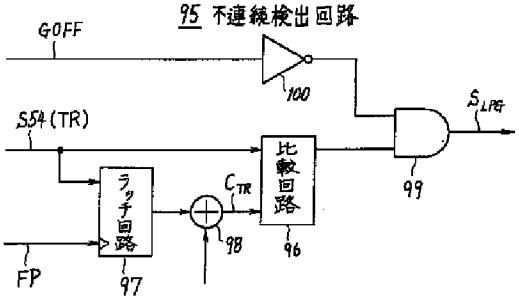


図13　他の実施例

【図9】

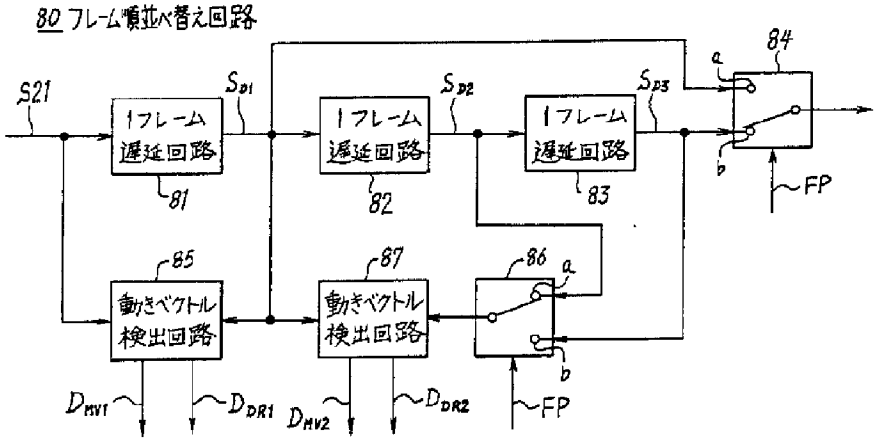


図9　フレーム順並べ替え回路

【図14】

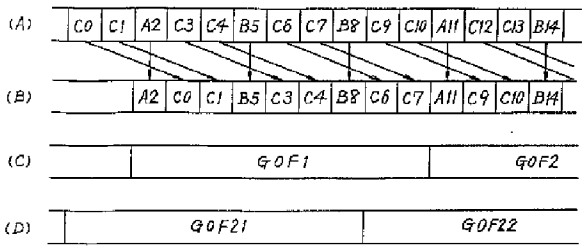


図14　他の実施例の記録順序

【図15】

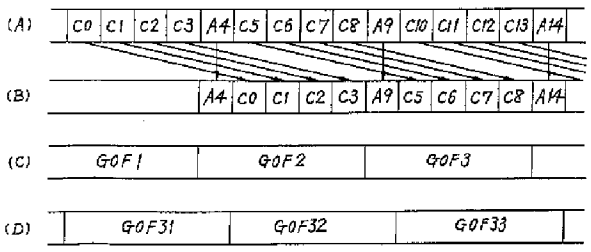


図15　他の実施例の記録順序

【図16】

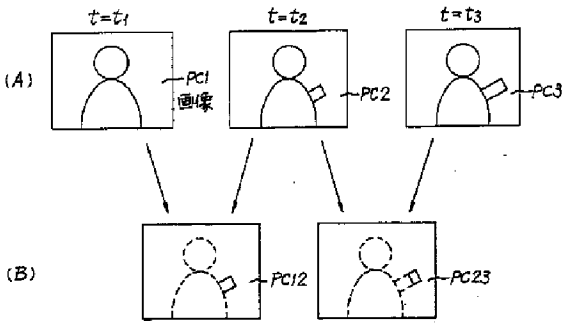


図16　フレーム内／フレーム間符号化処理

【図18】

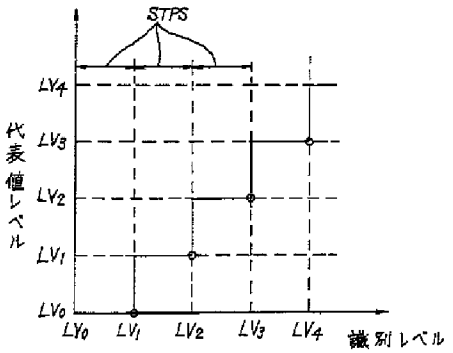


図18　量子化ステップ

【図10】

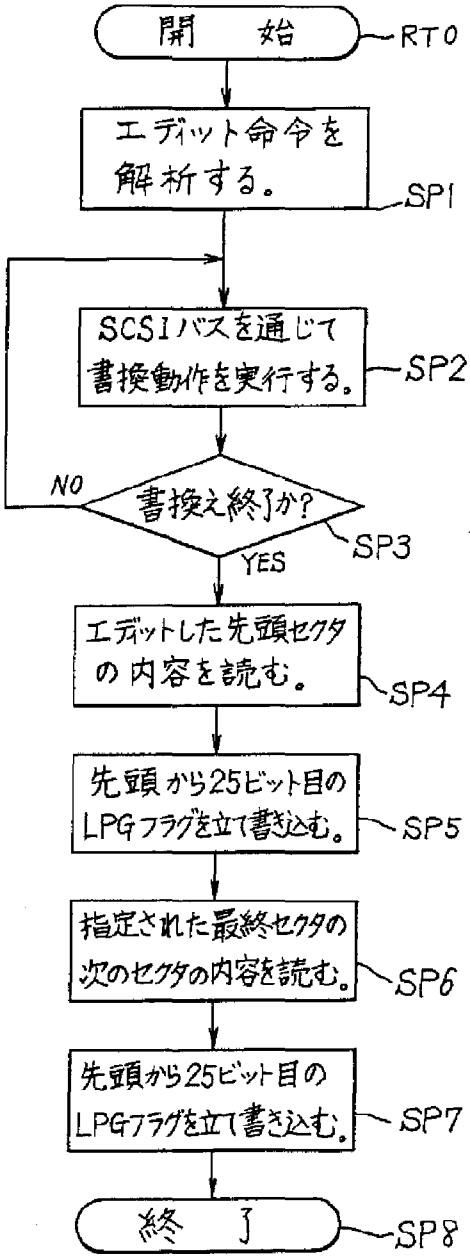


図１０　エディット処理手順

【図17】

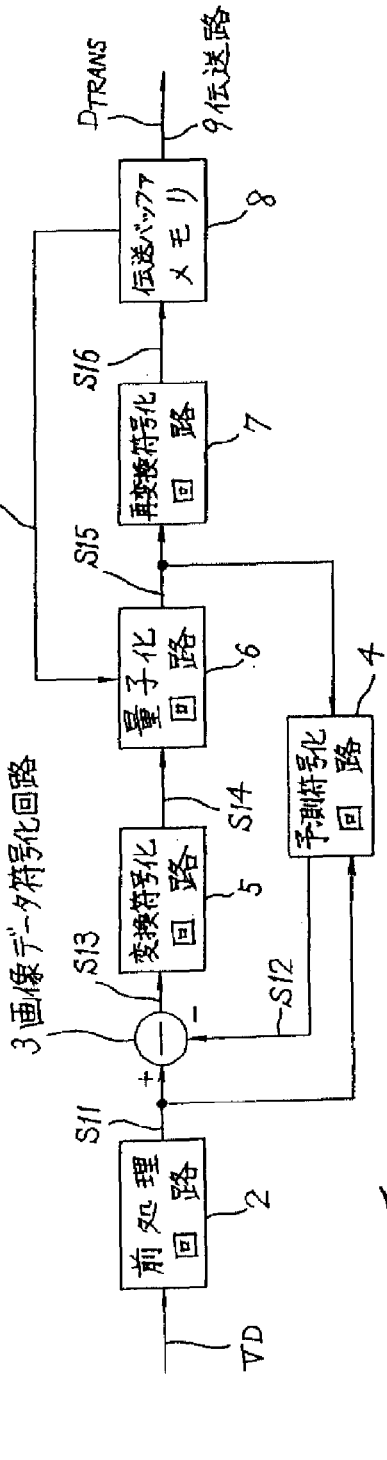


図17　動画符号化データ発生装置の概略構成

【図19】

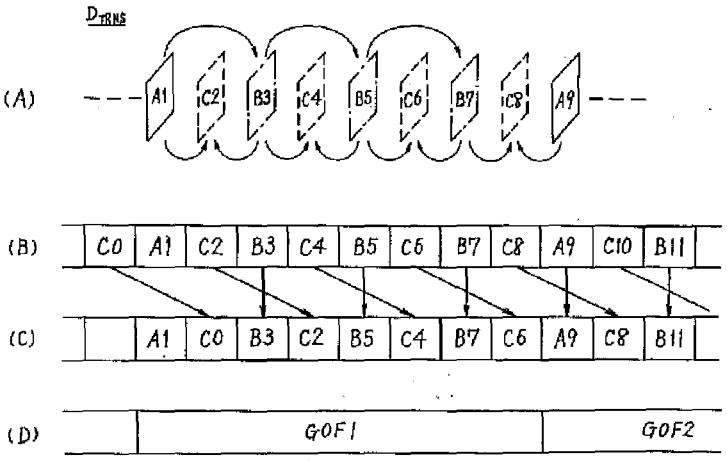


図19　従来の動画符号化データ記録順序

IN THE UNITED STATES PATENT & TRADEMARK OFFICE

Applicant:	HANNUKSELA	Examiner:	Lee, S.
Serial No.:	11/338,934	Group Art Unit:	2611
Filed:	January 25, 2006	Docket No.:	TPO.004.C2
Confirmation No.:	6610	Customer No.:	10888
Title:	GROUPING OF IMAGE FRAMES IN VIDEO CODING		

SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT
UNDER 37 C.F.R. § 1.97(d)

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

Applicant respectfully submits the items of information on the enclosed Form 1449 for the attention of the Examiner in the above-identified application.

Each item of information contained in the information disclosure statement was first cited in a communication from a foreign patent office in a counterpart foreign application not more than three months prior to the filing of this information disclosure statement. Please charge Deposit Account 50-3581 (TPO.004.C2) in the amount of $180.00 under 37 C.F.R. § 1.17(p) for consideration of the items.

In accordance with 37 C.F.R. § 1.98(a)(2), only a copy of each foreign document or non-U.S. patent/application listed on the enclosed Form 1449 is provided.

Please note that any notations or markings on the attached documents do not reflect particular relevance, or lack thereof, to the present application, nor were they necessarily made by anyone affiliated with the prosecution of the present application.

No representation is made that a reference is "prior art" within the meaning of 35 U.S.C. §§ 102 and 103. In addition, Applicant does not represent that a reference has been thoroughly reviewed or that any relevance of any portion of a reference is intended, and reserves the right to establish otherwise under 37 C.F.R. § 1.131 or others.

Consideration of the items listed is respectfully requested. According to MPEP § 609, Applicant requests that the Examiner return a copy of the attached Form 1449, marked as being considered and initialed by the Examiner, to the undersigned with the next official communication.

Respectfully submitted,

HOLLINGSWORTH & FUNK, LLC
8500 Normandale Lake Blvd., Suite 320
Minneapolis, MN 55437
952.232.6907

By: /Erin Nichols Matkaiti/
Erin Nichols Matkaiti
Reg. No. 57,125

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
11/338,934	01/25/2006	Miska Hannuksela	TPO.004.C2	6610

10888 7590 08/29/2011

Hollingsworth & Funk
8500 Normandale Lake Blvd., Suite 320
Minneapolis, MN 55437

EXAMINER	
LEE, SIU M	

ART UNIT	PAPER NUMBER
2611	

NOTIFICATION DATE	DELIVERY MODE
08/29/2011	ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

tdotter@hfiplaw.com
roswood@hfiplaw.com

Supplemental Notice of Allowability	Application No.	Applicant(s)	
	11/338,934	HANNUKSELA, MISKA	
	Examiner	Art Unit	
	SIU LEE	2611	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address--

All claims being allowable, PROSECUTION ON THE MERITS IS (OR REMAINS) CLOSED in this application. If not included herewith (or previously mailed), a Notice of Allowance (PTOL-85) or other appropriate communication will be mailed in due course. **THIS NOTICE OF ALLOWABILITY IS NOT A GRANT OF PATENT RIGHTS.** This application is subject to withdrawal from issue at the initiative of the Office or upon petition by the applicant. See 37 CFR 1.313 and MPEP 1308.

1. ☒ This communication is responsive to *IDS filed on 8/19/2011*.

2. ☐ An election was made by the applicant in response to a restriction requirement set forth during the interview on _____; the restriction requirement and election have been incorporated into this action.

3. ☒ The allowed claim(s) is/are *1, 2, 4-6, 8-11, 13-14*.

4. ☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
   a) ☒ All b) ☐ Some* c) ☐ None of the:
      1. ☒ Certified copies of the priority documents have been received.
      2. ☐ Certified copies of the priority documents have been received in Application No. _____ .
      3. ☐ Copies of the certified copies of the priority documents have been received in this national stage application from the International Bureau (PCT Rule 17.2(a)).

   * Certified copies not received: _____.

Applicant has THREE MONTHS FROM THE "MAILING DATE" of this communication to file a reply complying with the requirements noted below. Failure to timely comply will result in ABANDONMENT of this application.
THIS THREE-MONTH PERIOD IS NOT EXTENDABLE.

5. ☐ A SUBSTITUTE OATH OR DECLARATION must be submitted. Note the attached EXAMINER'S AMENDMENT or NOTICE OF INFORMAL PATENT APPLICATION (PTO-152) which gives reason(s) why the oath or declaration is deficient.

6. ☐ CORRECTED DRAWINGS (as "replacement sheets") must be submitted.
   (a) ☐ including changes required by the Notice of Draftsperson's Patent Drawing Review (PTO-948) attached
      1) ☐ hereto or 2) ☐ to Paper No./Mail Date _____.
   (b) ☐ including changes required by the attached Examiner's Amendment / Comment or in the Office action of Paper No./Mail Date _____.

Identifying indicia such as the application number (see 37 CFR 1.84(c)) should be written on the drawings in the front (not the back) of each sheet. Replacement sheet(s) should be labeled as such in the header according to 37 CFR 1.121(d).

7. ☐ DEPOSIT OF and/or INFORMATION about the deposit of BIOLOGICAL MATERIAL must be submitted. Note the attached Examiner's comment regarding REQUIREMENT FOR THE DEPOSIT OF BIOLOGICAL MATERIAL.

Attachment(s)

1. ☐ Notice of References Cited (PTO-892)
2. ☐ Notice of Draftperson's Patent Drawing Review (PTO-948)
3. ☒ Information Disclosure Statements (PTO/SB/08), Paper No./Mail Date 8/19/2011
4. ☐ Examiner's Comment Regarding Requirement for Deposit of Biological Material
5. ☐ Notice of Informal Patent Application
6. ☐ Interview Summary (PTO-413), Paper No./Mail Date _____ .
7. ☐ Examiner's Amendment/Comment
8. ☐ Examiner's Statement of Reasons for Allowance
9. ☐ Other _____.

Siu M Lee
Primary Examiner
Art Unit: 2611

FORM 1449* INFORMATION DISCLOSURE STATEMENT IN AN APPLICATION (Use several sheets if necessary)	Docket Number: TPO.004.C2	Application Number: 11/338,934
	Applicant: HANNUKSELA	
	Filing Date: 1/25/2006	Group Art Unit: 2611

U.S. PATENT DOCUMENTS

EXAMINER INITIAL	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE

FOREIGN PATENT DOCUMENTS

	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	TRANSLATION	
						YES	NO
/SL/	JP 8265694	10/11/1996	JP			Abstract	

OTHER DOCUMENTS (Including Author, Title, Date, Pertinent Pages, Etc.)

/SL/	8/2/2011	Notice of Allowance dated 8/2/2011 from Japanese Application No. 2006-120296, 3 pages.

EXAMINER /Siu Lee/	DATE CONSIDERED 08/22/2011

EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609; draw line through citation if not in conformance and not considered. Include copy of this form for next communication to the Applicant.

*Substitute Disclosure Statement Form (PTO-1449)

PTO/SB/47 (03-09)
Approved for use through 03/31/2012. OMB 0651-0016
U.S. Patent and Trademark Office; U. S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

"FEE ADDRESS" INDICATION FORM

Address to:
Mail Stop M Correspondence
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

- OR -

Fax to:
571-273-6500

INSTRUCTIONS: The issue fee must have been paid for application(s) listed on this form. In addition, only an address represented by a Customer Number can be established as the fee address for maintenance fee purposes (hereafter, fee address). A fee address should be established when correspondence related to maintenance fees should be mailed to a different address than the correspondence address for the application. **When to check the first box below**: If you have a Customer Number to represent the fee address. **When to check the second box below**: If you have no Customer Number representing the desired fee address, in which case a completed Request for Customer Number (PTO/SB/125) must be attached to this form. For more information on Customer Numbers, see the Manual of Patent Examining Procedure (MPEP) § 403.

For the following listed application(s), please recognize as the "Fee Address" under the provisions of 37 CFR 1.363 the address associated with:

☑ Customer Number: 00197

OR

☐ The attached Request for Customer Number (PTO/SB/125) form.

PATENT NUMBER (if known)	APPLICATION NUMBER
	11/338,934

Completed by (check one):

☐ Applicant/Inventor — /Erin Nichols Matkaiti/ (Signature)

☑ Attorney or Agent of record 57,125 (Reg. No.) — Erin Nichols Matkaiti (Typed or printed name)

☐ Assignee of record of the entire interest. See 37 CFR 3.71. Statement under 37 CFR 3.73(b) is enclosed. (Form PTO/SB/96) — 952-854-2700 (Requester's telephone number)

☐ Assignee recorded at Reel _____ Frame_____ — September 20, 2011 (Date)

NOTE: Signatures of all the inventors or assignees of record of the entire interest or their representative(s) are required. Submit multiple forms if more that one signature is required, see below*.

☑ * Total of 1 forms are submitted.

This collection of information is required by 37 CFR 1.363. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1. 11 and 1.14. This collection is estimated to take 5 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313- 1450. DO NOT SEND COMPLETED FORMS TO THIS ADDRESS.
SEND TO: Mail Stop M Correspondence, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.
If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

Privacy Act Statement

The **Privacy Act of 1974 (P.L. 93-579)** requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether disclosure of these records is required by the Freedom of Information Act.
2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.
3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.
4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).
5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.
6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).
7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (*i.e.*, GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.
8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspection or an issued patent.
9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

Electronic Patent Application Fee Transmittal

Application Number:	11338934
Filing Date:	25-Jan-2006
Title of Invention:	GROUPING OF IMAGE FRAMES IN VIDEO CODING
First Named Inventor/Applicant Name:	Miska Hannuksela
Filer:	Erin M. Nichols/Rennae Johnson
Attorney Docket Number:	TPO.004.C2

Filed as Large Entity

Utility under 35 USC 111(a) Filing Fees

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Basic Filing:				
Pages:				
Claims:				
Miscellaneous-Filing:				
Petition:				
Patent-Appeals-and-Interference:				
Post-Allowance-and-Post-Issuance:				
Utility Appl issue fee	1501	1	1510	1510
Publ. Fee- early, voluntary, or normal	1504	1	300	300

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Extension-of-Time:				
Miscellaneous:				
			Total in USD ($)	**1810**

Electronic Acknowledgement Receipt

EFS ID:	10986678
Application Number:	11338934
International Application Number:	
Confirmation Number:	6610
Title of Invention:	GROUPING OF IMAGE FRAMES IN VIDEO CODING
First Named Inventor/Applicant Name:	Miska Hannuksela
Customer Number:	10888
Filer:	Erin M. Nichols
Filer Authorized By:	
Attorney Docket Number:	TPO.004.C2
Receipt Date:	20-SEP-2011
Filing Date:	25-JAN-2006
Time Stamp:	17:29:02
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	yes
Payment Type	Deposit Account
Payment was successfully received in RAM	$1810
RAM confirmation Number	5183
Deposit Account	503581
Authorized User	

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)

1	Issue Fee Payment (PTO-85B)	IssueFee-TPO004C2.pdf	82510 a1cdcacf7a6b4c54f52792a92b7ce6fe4d8d6b8a	no	2
Warnings:					
Information:					
2	Change of Address	FeeAddress-TPO004C2.pdf	65824 5ce644d57d545fa0e1f17f976f87d9a79298e242	no	2
Warnings:					
Information:					
3	Fee Worksheet (SB06)	fee-info.pdf	32034 0c9fa799ec5a84da31c2d64c3cd412a09979ef6f	no	2
Warnings:					
Information:					
		Total Files Size (in bytes):	180368		

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

PART B - FEE(S) TRANSMITTAL

Complete and send this form, together with applicable fee(s), to: Mail **Mail Stop ISSUE FEE**
Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450
or Fax (571) 273-2885

INSTRUCTIONS: This form should be used for transmitting the ISSUE FEE and PUBLICATION FEE (if required). Blocks 1 through 5 should be completed where appropriate. All further correspondence including the Patent, advance orders and notification of maintenance fees will be mailed to the current correspondence address as indicated unless corrected below or directed otherwise in Block 1, by (a) specifying a new correspondence address; and/or (b) indicating a separate "FEE ADDRESS" for maintenance fee notifications.

CURRENT CORRESPONDENCE ADDRESS (Note: Use Block 1 for any change of address)

HOLLINGSWORTH & FUNK, LLC
8500 Normandale Lake Blvd.
Suite 320
Minneapolis, MN 55437

Note: A certificate of mailing can only be used for domestic mailings of the Fee(s) Transmittal. This certificate cannot be used for any other accompanying papers. Each additional paper, such as an assignment or formal drawing, must have its own certificate of mailing or transmission.

Certificate of Mailing or Transmission

I hereby certify that this Fee(s) Transmittal is being deposited with the United States Postal Service with sufficient postage for first class mail in an envelope addressed to the Mail Stop ISSUE FEE address above, or being facsimile transmitted to the USPTO (571) 273-2885, on the date indicated below.

(Depositor's name)
(Signature)
(Date)

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
11/338,934	01/25/2006	Miska Hannuksela	TPO.004.C2	6610

TITLE OF INVENTION:

APPLN. TYPE	SMALL ENTITY	ISSUE FEE	PUBLICATION FEE	TOTAL FEE(S) DUE	DATE DUE
nonprovisional	NO	$1510	$300	$1810	11/18/2011

EXAMINER	ART UNIT	CLASS-SUBCLASS

1. Change of correspondence address or indication of "Fee Address" (37 CFR 1.363).

☐ Change of correspondence address (or Change of Correspondence Address form PTO/SB/122) attached.

☑ "Fee Address" indication (or "Fee Address" Indication form PTO/SB/47; Rev 03-02 or more recent) attached. **Use of a Customer Number is required.**

2. For printing on the patent front page, list

(1) the names of up to 3 registered patent attorneys or agents OR, alternatively,

(2) the name of a single firm (having as a member a registered attorney or agent) and the names of up to 2 registered patent attorneys or agents. If no name is listed, no name will be printed.

1 _
2 HOLLINGSWORTH & FUNK, LLC
3 _

3. ASSIGNEE NAME AND RESIDENCE DATA TO BE PRINTED ON THE PATENT (print or type)

PLEASE NOTE: Unless an assignee is identified below, no assignee data will appear on the patent. If an assignee is identified below, the document has been filed for recordation as set forth in 37 CFR 3.11. Completion of this form is NOT a substitute for filing an assignment.

(A) NAME OF ASSIGNEE	(B) RESIDENCE: (CITY and STATE OR COUNTRY)
Nokia Corporation	Espoo, Finland

Please check the appropriate assignee category or categories (will not be printed on the patent) : ☐ Individual ☑ Corporation or other private group entity ☐ Government

4a. The following fee(s) are enclosed:

☑ Issue Fee
☑ Publication Fee (No small entity discount permitted)
☐ Advance Order - # of Copies _

4b. Payment of Fee(s):

☐ A check in the amount of the fee(s) is enclosed.
☐ Payment by credit card. Form PTO-2038 is attached.
☑ The Director is hereby authorized by charge the required fee(s), or credit any overpayment, to Deposit Account Number 50-3581 .

5. **Change in Entity Status** (from status indicated above)

☐ a. Applicant claims SMALL ENTITY status. See 37 CFR 1.27.
☐ b. Applicant is no longer claiming SMALL ENTITY status. See 37 CFR 1.27(g)(2).

The Director of the USPTO is requested to apply the Issue Fee and Publication Fee (if any) or to re-apply any previously paid issue fee to the application identified above.
NOTE: The Issue Fee and Publication Fee (if required) will not be accepted from anyone other than the applicant; a registered attorney or agent; or the assignee or other party in interest as shown by the records of the United States Patent and Trademark Office.

Authorized Signature /Erin Nichols Matkaiti/ Date September 20, 2011

Typed or printed name Erin Nichols Matkaiti Registration No. 57,125

This collection of information is required by 37 CFR 1.311. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, Virginia 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, Virginia 22313-1450.

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether disclosure of these records is required by the Freedom of Information Act.
2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.
3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.
4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).
5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.
6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).
7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.
8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspection or an issued patent.
9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

FORM 1449* INFORMATION DISCLOSURE STATEMENT IN AN APPLICATION (Use several sheets if necessary)	Docket Number: TPO.004.C2	Application Number: 11/338,934
	Applicant: HANNUKSELA	
	Filing Date: Jan. 25, 2006	Group Art Unit: 2613

U.S. PATENT DOCUMENTS

EXAMINER INITIAL	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
/SL/	~~6,26,158~~ 6,266,158	7/24/2001	Hata et al.			

Change(s) applied to document, /S.R.R./ 9/15/2011

FOREIGN PATENT DOCUMENTS

	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	TRANSLATION	
						YES	NO
/SL/	10-271511	10/09/1998	JP				X

OTHER DOCUMENTS (Including Author, Title, Date, Pertinent Pages, Etc.)

EXAMINER /Siu Lee/	DATE CONSIDERED 05/20/2010

EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609; draw line through citation if not in conformance and not considered. Include copy of this form for next communication to the Applicant.

*Substitute Disclosure Statement Form (PTO-1449) Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	ISSUE DATE	PATENT NO.	ATTORNEY DOCKET NO.	CONFIRMATION NO.
11/338,934	11/01/2011	8050321	TPO.004.C2	6610

10888 7590 10/12/2011

Hollingsworth & Funk
8500 Normandale Lake Blvd., Suite 320
Minneapolis, MN 55437

ISSUE NOTIFICATION

The projected patent number and issue date are specified above.

Determination of Patent Term Adjustment under 35 U.S.C. 154 (b)
(application filed on or after May 29, 2000)

The Patent Term Adjustment is 1632 day(s). Any patent to issue from the above-identified application will include an indication of the adjustment on the front page.

If a Continued Prosecution Application (CPA) was filed in the above-identified application, the filing date that determines Patent Term Adjustment is the filing date of the most recent CPA.

Applicant will be able to obtain more detailed information by accessing the Patent Application Information Retrieval (PAIR) WEB site (http://pair.uspto.gov).

Any questions regarding the Patent Term Extension or Adjustment determination should be directed to the Office of Patent Legal Administration at (571)-272-7702. Questions relating to issue and publication fee payments should be directed to the Application Assistance Unit (AAU) of the Office of Data Management (ODM) at (571)-272-4200.

APPLICANT(s) (Please see PAIR WEB site http://pair.uspto.gov for additional applicants):

Miska Hannuksela, Tampere, FINLAND;