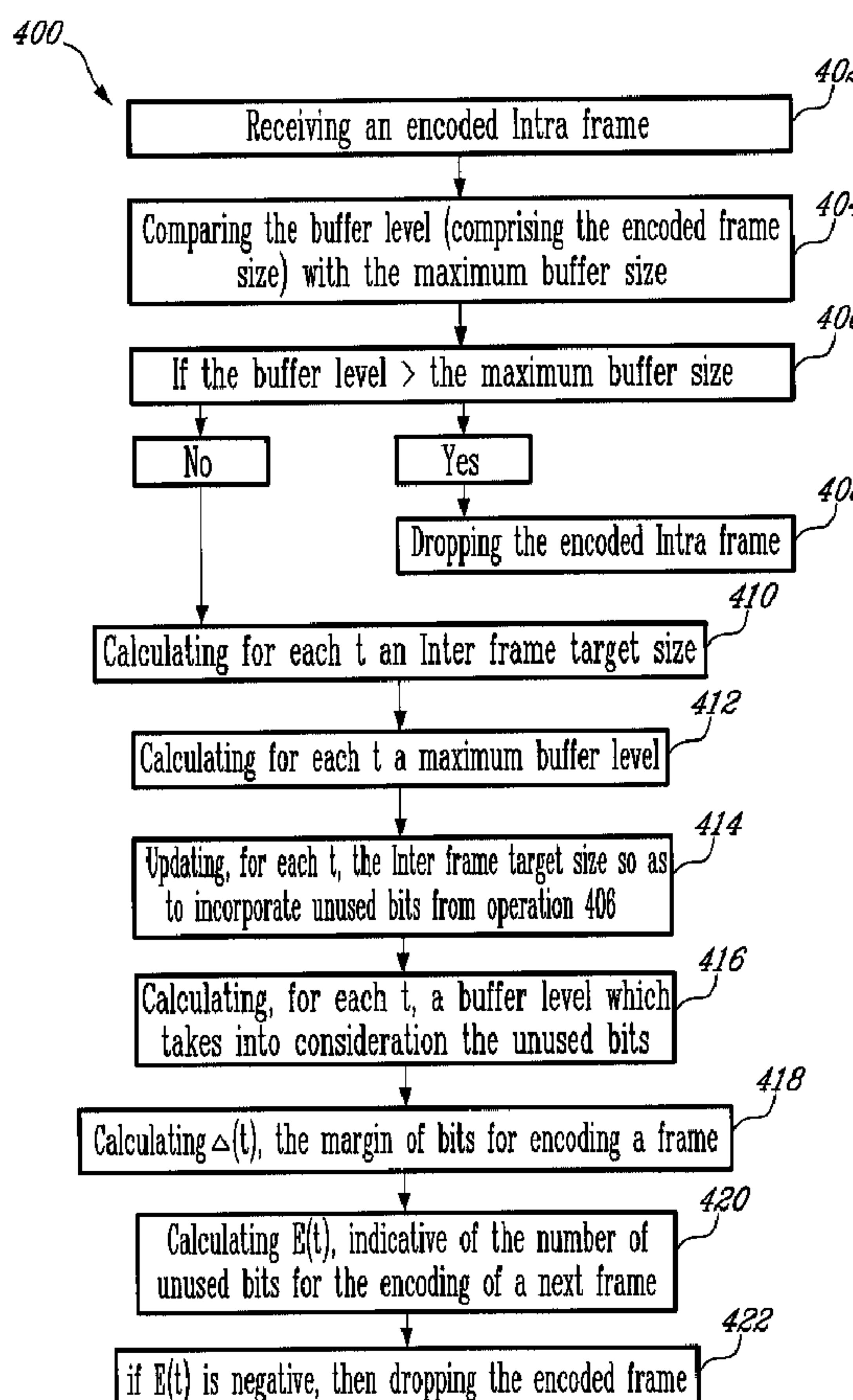




(86) Date de dépôt PCT/PCT Filing Date: 2007/12/11
 (87) Date publication PCT/PCT Publication Date: 2008/06/19
 (45) Date de délivrance/Issue Date: 2018/12/04
 (85) Entrée phase nationale/National Entry: 2009/06/11
 (86) N° demande PCT/PCT Application No.: CA 2007/002242
 (87) N° publication PCT/PCT Publication No.: 2008/070987
 (30) Priorité/Priority: 2006/12/12 (US60/874,250)

(51) CI.Int./Int.Cl. H04N 19/146 (2014.01),
 H04N 19/172 (2014.01), H04N 19/61 (2014.01)
 (72) Inventeur/Inventor:
 COULOMBE, STEPHANE, CA
 (73) Propriétaire/Owner:
 VANTRIX CORPORATION, CA
 (74) Agent: DONNELLY, VICTORIA

(54) Titre : COMMANDÉ DE DEBIT VIDEO AMELIORE POUR NORMES DE CODAGE VIDEO
 (54) Title: AN IMPROVED VIDEO RATE CONTROL FOR VIDEO CODING STANDARDS



(57) Abrégé/Abstract:

A method and device for improving rate controlling in video coding of sequences including a series of Inter frames separated by Intra frames, when a decoding delay is considered, comprise for each Inter frame of the series: computing a target frame size,

(57) Abrégé(suite)/Abstract(continued):

computing a maximum buffer level related to a position of each Inter frame relative to a previous Intra frame and an upcoming Intra frame, and optimizing a transmission buffer level in response to the computed target frame size and the computed maximum buffer level.

ABSTRACT OF THE DISCLOSURE

A method and device for improving rate controlling in video coding of sequences including a series of Inter frames separated by Intra frames, when a decoding delay is considered, comprise for each Inter frame of the series: computing a target frame size, computing a maximum buffer level related to a position of each Inter frame relative to a previous Intra frame and an upcoming Intra frame, and optimizing a transmission buffer level in response to the computed target frame size and the computed maximum buffer level.

TITLE OF THE INVENTION

AN IMPROVED VIDEO RATE CONTROL FOR VIDEO
CODING STANDARDS

FIELD OF THE INVENTION

[001] The present invention generally relates to video coding. More specifically, the present invention is concerned with video rate controlling in video coding, for example but not exclusively for video coding standards such as the *H.263* and *MPEG-4* (Moving Picture Experts Group-4) standards.

BACKGROUND OF THE INVENTION

[002] The past few years have witnessed a great popularity of digital and online videos and their applications. With the emergence of fast communication technologies and multimedia applications, digital video codecs are used in many areas and systems, such as in *DVDs* (Digital Video Disc) employing the *MPEG-2* (Moving Picture Experts Group-2) format, in *VCDs 2* (Video Compact Disc) employing the *MPEG-1* (Moving Picture Experts Group-1) format, in emerging satellite and terrestrial broadcast systems, and on the Internet.

[003] More specifically, this popularity of video applications allowed for interesting developments in video codecs, which compress and decompress video data. In video data compression, a balance is kept between the video quality and the compression rate, i.e. the necessary transmitted quantity of data, in other words, the bitrate needed to represent a video.

[004] In addition, the complexity of encoding and decoding algorithms, robustness to data losses and errors, the state of the art of compression

algorithm design, end-to-end delay in a video conference application for example, etc., are also considered.

[005] A plurality of video coding standards exist, each of them is specially designed for a particular type of application. For example, the *H.263* standard, published by the *ITU* (International Telecommunications Union) is a video coding and compression standard for low bitrates, such as in the range of 40-128 kbps (kilobits per second). More specifically, this standard supports video coding in video-conferencing and video-telephony applications.

[006] The *H.263* standard specifies the format and content of the encoded stream of data, therefore, it sets the requirements for the encoder and decoder to meet, without specifically providing a design or structure of an encoder and decoder themselves. Similar principles apply to other video standards such as **MPEG-4**.

[007] In video compression, each picture is represented by typically two kinds of pictures, commonly referred to as frames, i.e. the **Intra** frames and **Inter** frames. Furthermore, the **Inter** frames are separated into two categories, i.e. the **P-frames** (Predictive frames) and **B-frames** (Bi-predictive or Bi-directional frames). The **Intra** frames represent a whole picture, therefore they are bandwidth consuming since the content of the whole picture must be encoded. In order to compress and therefore save bandwidth, only differences between whole pictures (or **Intra** frames) are encoded and then transmitted. Those differences are represented by the **P-frames** and the **B-frames**. For example, the background between two consecutive pictures usually do not change, therefore, the background does not need to be encoded again. The **B-frames** are bi-directional and thus perform a bi-directional prediction, i.e. a prediction with the previous and next pictures.

[008] Furthermore, when compressing videos, a picture is divided into macroblocks for processing purposes. Indeed, processing is applied macroblock by macroblock. Each macroblock generally represents a block of 16 by 16 pixels.

[009] A video encoder generally includes a motion estimation module, a motion compensation module, a *DCT* (Discrete Cosine Transforms) module, and a quantizing module.

[0010] The motion estimation module allows for predicting which areas of a previous frame have been moved into the current frame so that those areas do not need to be re-encoded.

[0011] The motion compensation module allows for compensating for the movement of the areas from the previous frame into the current frame.

[0012] *DCT* are generally used for transforming a block of pixels into “spatial frequency coefficients”. They operate on a two-dimensional block of pixels, such as a macroblock. Since *DCT* are efficient at compacting the energy (or information) of pictures, generally a few *DCT* coefficients are sufficient for recreating the original picture.

[0013] Also, a quantizing module is provided for quantizing the *DCT* coefficients. For example, the quantizing module sets the near zero *DCT* coefficients to zero and quantizes the remaining non-zero *DCT* coefficients.

[0014] One of the limitations in video coding comes from the capacity of a channel. Indeed, communication channels are limited by the number of bits that they can transmit per second. In many channels, the bitrate is constant, such

as in *ISDN* (Integrated Services Digital Networks), *POTS* (Plain Old Telephone Service), wireless channels, etc.

[0015] However, depending on the efficiency of the algorithms used to compress the videos and the motion complexity of those videos, the bit budget and the bitrate needed for encoding and transmitting the encoded videos may vary or increase. Therefore a rate control is used to adjust the bitrate required for encoding videos of various complexity to the bitrate of the channel used to transmit those encoded videos.

[0016] The current rate control algorithm used in the *H.263* standard is called the *TMN8* (Test Model Near-Term version 8). Generally stated, this rate control algorithm ensures that only an average bitrate is met.

[0017] The paper entitled “*Rate Control in DCT Video Coding for Low-Delay Communications*”, by Jordi Ribas-Corbera, 1999, hereinafter referred to as Reference 1, discloses an algorithm used by the rate control *TMN8* to ensure that the target average bitrate, related to a target frame size, is met by each frame. More specifically, the *TMN8* rate control algorithm computes some image statistics to determine some proper *QP* (Quantization Parameter) values for each macroblock and update them within each Inter frame so as to meet the target frame size. Unfortunately, this control is very approximate and often the resulting frame size can be significantly over or under the target frame size. For Intra frames, a fixed *QP* is used for the whole video sequence regardless of the characteristics of the video sequence. Having no control over the size of Intra frames is generally a factor leading to exceeding the desired bitrate.

[0018] Furthermore, the rate control *TMN8* cannot control both an average target bitrate and a maximum bitrate. Indeed, the *TMN8* rate control algorithm used in the *H.263* video coding standard only uses an average bitrate

parameter. However, in many video applications, a maximum bitrate should also be considered in addition to the average bitrate.

[0019] *TMN8* cannot guarantee to not exceed a given target bitrate, since the encoder has no control over the Intra frame sizes and no sufficient control over the Inter frame sizes. When the given target bitrate is exceeded, the encoder will skip a certain number of frames so as to compensate for the overflow. However, by so doing, the quality of the communication and the videos is altered.

[0020] Therefore, there is a need for overcoming the above discussed problems, related to the limitations of the current rate control in video coding standards, such as the *H.263* standard. Accordingly, a method and system for improving the rate control in video coding standard are sought.

OBJECTS OF THE INVENTION

[0021] An object of the present invention is therefore to provide rate control in video coding, for example in video coding standards such as the *H.263* and *MPEG-4* standards, so as to enhance the quality of the videos by ensuring that an average bitrate and a maximum bit rate are both respected.

SUMMARY OF THE INVENTION

[0022] According to one aspect of the invention, there is provided a method for rate controlling in video coding of sequences of pictures, including a series of Inter frames separated by Intra frames, the method comprising:

(a) for an Inter frame in the series, determining a maximum buffer level of a transmission buffer as a function of a position of the Inter frame relative to a

previous Intra frame and an upcoming Intra frame in the sequence of pictures;

- (b) determining a target frame size measured in bits for the Inter frame leading to a transmission buffer level not exceeding the maximum buffer level;
- (c) encoding the Inter frame into an encoded frame according to the target frame size, the encoded frame having an encoded frame size;
- (d) determining the transmission buffer level in response to the encoded frame; and
- (e) dropping the encoded frame if it would cause the maximum buffer level to be exceeded.

[0023] In the method described above, the step (e) further including restoring the transmission buffer level to a previous value, provided the encoded frame was dropped.

[0024] In the method described above, the step (a) comprises determining the maximum buffer level, which ensures that a transmission buffer can accommodate the upcoming Intra frame.

[0025] In the method described above, the step (a) comprises determining the maximum buffer level close to a maximum buffer size when an Inter frame to be encoded is near the previous Intra frame, and close to zero when the Inter frame is near the upcoming Intra frame.

[0026] In the method described above, the step (c) comprises:

determining a difference between the maximum buffer level and the transmission buffer level; and

determining a margin of error between the target frame size and the maximum buffer size for encoding the Inter frames and the Intra frames based on the determined difference while keeping said margin of error.

[0027] In the method described above, the step (a) comprises determining the maximum buffer level as follows:

$$\begin{aligned} \text{BufferLevel}_{MAX}(t) &= R_{MAX} && \text{if } t=0,1,\dots,L \\ \text{BufferLevel}_{MAX}(t) &= \frac{R_{MAX}(FT-F-L)}{F(T-L-1)} + t \frac{R_{MAX}(1-F)}{F(T-F)} && \text{if } t = L+1, \dots, T-1 \end{aligned}$$

wherein $\text{BufferLevel}_{MAX}(t)$ is the maximum buffer level, T is a period between two Intra frames, F is a frame rate, R_{MAX} is a maximum bitrate, L is a number of frames over which the transmission buffer level is optimized, and “ t ” is time.

In the method described above, the determining the target frame size comprises:

- (a-1) determining a number of bits unused in encoding an Intra frame; and
- (a-2) distributing the number of unused bits over the target frame size for N Inter frames following the Intra frame,

wherein N is a number of Inter frames over which a number of unused bits are distributed.

[0028] In the method described above:

the step (a-1) comprises determining a difference between the encoded Intra frame size and a corresponding Intra frame target size; or

the step (a-2) comprises optimizing the transmission buffer level over a number L of Inter frames, where $0 \leq N \leq L$; or

the step (a-2) comprises distributing the number of unused bits substantially equally.

[0029] In the method described above, the step (b) comprises determining the target frame size as follows:

$$B_p(t) = \begin{cases} B_p^I(t) & \text{if } t = 1, 2, 3, \dots, L \\ \frac{TR_{MAX} - FB_I - F \sum_{k=1}^L B_p^I(k)}{F(T - L - 1)} & \text{if } t = L, L+1, \dots, T-1 \end{cases}$$

where $B_p(t)$ is the target frame size of an Inter frame at time t , T is a period between two Intra frames, $B_p^I(t)$ is an Inter frame target size at time t , which is selected so that not to lead to overflowing the transmission buffer, B_I is an Intra frame size, F is a frame rate, k is an index in a summation function, L is a number of frames over which the transmission buffer level is optimized, and R_{MAX} is a maximum bitrate.

[0030] In the method described above, the step (b) comprises determining the target frame size as follows:

$$B_p(t) = \begin{cases} \frac{R_{MAX}}{F} + \frac{(R_{MAX} - B_I)}{N} & \text{if } 0 < t \leq N \\ \frac{R_{MAX}}{F} & \text{if } N < t \leq L \end{cases}$$

$$B_p(t) = \frac{R_{MAX}(T-F-L)}{F(T-L-1)} \quad \text{if } t = L, L+1, \dots, T-1$$

where $B_p(t)$ is the target frame size of the Inter frame at time t , B_I is an Intra frame size, T is a period between two Intra frames, B_p^I is an Inter frame target size equal to R_{MAX}/F , F is a frame rate, L is a number of frames over which the transmission buffer level is optimized, and R_{MAX} is a maximum bitrate.

[0031] The method described further comprises:

- (i) dropping the Inter frames when a size of the Inter frame after encoding exceeds the maximum buffer level; or
- (ii) storing respective encoder state variables before encoding the Intra frames or the Inter frames; or
- (iii) restoring the respective encoder state variables after an encoded Inter frame or an encoded Intra frame has been dropped prior to encoding a next Inter frame or a next Intra frame; or
- (iv) verifying that a size of the Inter frame or a size of the Intra frame after encoding is smaller than a given maximum limit, wherein the maximum limit is a maximum bitrate.

[0032] According to another aspect of the invention, there is provided a device for rate controlling in video coding of sequences pictures, including a series of Inter frames separated by Intra frames, the device comprising:

- (a) means for determining, for an Inter frame of the series, a maximum buffer level of a transmission buffer as a function of a position of the Inter frame

relative to a previous Intra frame and an upcoming Intra frame in the sequence of pictures;

(b) means for determining a target frame size measured in bits for the Inter frame leading to a transmission buffer level not exceeding the maximum buffer level;

(c) means for encoding the Inter frame into an encoded frame according to the target frame size, the encoded frame having an encoded frame size;

(d) means for determining the transmission buffer level in response to the encoded frame; and

(e) means for dropping the encoded frame if it would cause the maximum buffer level to be exceeded.

[0033] In the device described above, the means (e) comprises means for restoring the transmission buffer level to a previous value, provided the encoded frame was dropped.

[0034] In the device described above, the means (a) is configured:

to determine the maximum buffer level, which ensures that the transmission buffer can accommodate the upcoming Intra frame; or

to determine the maximum buffer level close to a maximum buffer size when an Inter frame to be encoded is near the previous Intra frame, and close to zero when the Inter frame is near the upcoming Intra frame; or

to determine the maximum buffer level as follows:

$$\begin{aligned}
 BufferLevel_{MAX}(t) &= R_{MAX} && \text{if } t=0,1,\dots,L \\
 BufferLevel_{MAX}(t) &= \frac{R_{MAX}(FT-F-L)}{F(T-L-1)} + t \frac{R_{MAX}(1-F)}{F(T-F)} && \text{if } t = L+1, \dots, T-1
 \end{aligned}$$

wherein $BufferLevel_{MAX}(t)$ is the maximum buffer level, T is a period between two Intra frames, F is a frame rate, R_{MAX} is a maximum bitrate, L is a number of frames over which the transmission buffer level is optimized, and “ t ” is time.

[0035] In the device described above, the means (c) is configured:

to determine a difference between the maximum buffer level and the transmission buffer level; and

to determine a margin of error between the target frame size and the maximum buffer size for encoding the Inter frames and the Intra frames based on the determined difference while keeping said margin of error.

[0036] In the device described above, the means (a) comprises:

(a-1) means for determining a number of bits unused in encoding an Intra frame; and

(a-2) means for distributing the number of unused bits over the target frame size for N Inter frames following the Intra frame,

wherein N is a number of Inter frames over which a number of unused bits are distributed.

[0037] In the device described above:

the means (a-1) is configured to determine a difference between the encoded Intra frame size and a corresponding Intra frame target size; or

the means (a-2) is configured to optimize the transmission buffer level over a number L of Inter frames, where $0 \leq N \leq L$; or

the means (a-2) is configured to distribute the number of unused bits substantially equally.

[0038] In the device described above:

the means (b) is configured to determine the target frame size as follows:

$$B_p(t) = B_p^I(t) \quad \text{if } t = 1, 2, 3, \dots, L$$

$$B_p(t) = \frac{TR_{MAX} - FB_I - F \sum_{k=1}^L B_p^I(k)}{F(T - L - 1)} \quad \text{if } t = L, L+1, \dots, T-1$$

where $B_p(t)$ is the target frame size of an Inter frame at time t, T is a period between two Intra frames, $B_p^I(t)$ is an Inter frame target size at time t, which is selected so that not to lead to overflowing the transmission buffer, B_I is an Intra frame size, F is a frame rate, k is an index in a summation function, L is a number of frames over which the transmission buffer level is optimized, and R_{MAX} is a maximum bitrate; or

the means (a) is configured to determine the target frame size as follows:

$$B_p(t) = \frac{R_{MAX}}{F} + \frac{(R_{MAX} - B_I)}{N} \quad \text{if } 0 < t \leq N$$

$$B_p(t) = B_p^I = \frac{R_{MAX}}{F} \quad \text{if } N < t \leq L$$

$$B_p(t) = \frac{R_{MAX}(T-F-L)}{F(T-L-1)} \quad \text{if } t = L, L+1, \dots, T-1$$

where $B_p(t)$ is the target frame size of the Inter frame at time t , B_i is an Intra frame size, T is a period between two Intra frames, B_p^I is an Inter frame target size equal to R_{MAX}/F , F is a frame rate, L is a number of frames over which the transmission buffer level is optimized, and R_{MAX} is a maximum bitrate.

[0039] The device described above has been further configured:

- (i) to drop the Inter frames when a size of the Inter frame after encoding exceeds the maximum buffer level; or
- (ii) to store respective encoder state variables before encoding the Intra frames or the Inter frames; or
- (iii) to restore the respective encoder state variables after an encoded Inter frame or an encoded Intra frame has been dropped prior to encoding a next Inter frame or a next Intra frame; or
- (iv) to verify that a size of the Inter frame or a size of the Intra frame after encoding is smaller than a given maximum limit, wherein the maximum limit is a maximum bitrate.

[0040] According to yet another aspect of the invention, there is provided an encoder for video coding of sequences of pictures, including a series of Inter frames separated by Intra frames, the encoder comprising:

a processor; and

the device for rate controlling in the video coding as described above.

[0041] The foregoing and other objects, advantages and features of the present invention will become more apparent upon reading of the following non-restrictive description of illustrative embodiments thereof, given by way of example only with reference to the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

[0042] In the appended drawings:

[0043] Figure 1 is a schematic diagram of a video communication system;

[0044] Figure 2 is a schematic block diagram of an encoder for video coding in the communication system of Figure 1;

[0045] Figure 3 is a schematic block diagram of a device for rate controlling in video coding in accordance with a non-restrictive illustrative embodiment of the present invention;

[0046] Figure 4 is a flow chart illustrating a method according to a non-restrictive illustrative embodiment of the present invention for rate control in the H.263 video coding standard;

[0047] Figure 5 is a flow chart illustrating an example of method for calculating an Inter frame target size;

[0048] Figure 6 is a flow chart illustrating another example of method for calculating an Inter frame target size;

[0049] Figure 7 is a flow chart illustrating an example of method for updating the calculated Inter frame target size;

[0050] Figure 8 is a flow chart illustrating another example of method for updating the calculated Inter frame target size; and

[0051] Figure 9 a flow chart illustrating a method according to another non-restrictive illustrative embodiment of the present invention for rate control in the H.263 video coding standard.

DETAILED DESCRIPTION

[0052] Generally stated, a device for improving the rate control according to the non-restrictive illustrative embodiment of the present invention allows for addressing the following requirements, which are not met with by the TMN8 rate control:

1. Being able to specify and meet with a target frame size for an Intra frame:
 - by so doing, the quality of the Intra frame and fluctuations in the bitrate caused by the presence of the Intra frame can be controlled; and
 - the QP (Quantization Parameters) are not fixed but computed based on the Intra frame characteristics and may change for each block of pixels of the Intra frame in order to meet the specified Intra frame target size.
2. Being able to specify and meet with a maximum bitrate in addition to an average bitrate:
 - furthermore, a validation on the encoded frame size is performed, after the frame has been encoded; if the encoded frame size is too large to be

acceptable, compared to the maximum bitrate, then the frame needs to be either re-encoded with a smaller target size or dropped;

- in the contrast, the *TMN8* rate control attempts to meet with only an average bitrate, by dropping frames for example; it cannot meet with a maximum bitrate.

3. Having a circular buffer to manage the maximum bitrate and another buffer for the average bitrate;

- in the contrast, the *TMN8* uses only one transmission buffer.

4. Estimating and calculating an Inter frame target size, which depends on the timing and encoding of an Intra frame: if an Intra frame has been recently encoded or if an Intra frame will soon be encoded;

- also, the estimates of the Inter frame target size takes into consideration various parameters, such as the size of the Intra frame, the period between two Intra frames, the number of bits under or over the average bitrate, etc.

- in the contrast, the *TMN8* uses a target size which is based on the transmission buffer occupancy and rate only; it never predicts upcoming Intra frames and typically skip many frames after the encoding of an Intra frame, therefore causing a degradation in the video quality.

5. Dropping frames in a proactive and preventive way:

- the device according to one non-restrictive illustrative embodiment of the present invention ensures that the maximum bitrate is never exceeded, and drops frames if required, before transmitting them; the decision to drop frames is based on a difference between what was estimated and what was achieved, i.e. what was actually encoded;
- in the contrast, the *TMN8* skips frames after having created and transmitted a frame that was too large and leading to exceeding the

target size; by so doing, harm has been already done to the quality of the videos; furthermore, the decision to drop frames is based on the level of fullness of the transmission buffer.

I. The system

[0053] First, a video communication system will be described so as to define a communication system in which video coding takes place. However, the communication system is not part of the present invention.

[0054] Turning to Figure 1, a communication system 10 will be described.

[0055] The communication system 10 comprises a video source 12, such as a video camera or web camera, for capturing videos. Those videos are then provided to a video encoder 14. The video encoder 14 is connected to a network 16. The network 16 can also include a communication link for transmitting the encoded videos to a decoder 18. Finally, the decoder 18 is connected to a video player 20. The video player 20 can be a computer, a television set, or any device which can display videos.

[0056] During a communication session between, for example, first and second parties, the videos captured by the video source 12 of the first party are received by the encoder 14, which then can perform video compressing and coding. Once the encoder 14 has finished encoding the video data, the encoded video is transmitted to the second party through the network 16 or simply a communication link. At the second party's end, the decoder 18 receives the encoded video and starts decoding it. Once the encoded video is decoded, the decoder 18 sends the decoded video data to the video player 20, which then can start playing the video received from the first party.

[0057] The encoder 14 is detailed in Figure 2. The encoder 14 includes one or a plurality of processes 30₁ to 30_N, which allows for performing different operations in video coding as will be described hereinafter. The encoder 14 further includes a motion estimation module 32, a motion compensation module 34, a DCT module 36, a quantizing module 38, and a buffer 40.

[0058] It should be noted that this encoder 14 can be implemented in different platforms, such as, for example, the spot xde® from Vantrix Corporation. The video source can be a video file or a video camera.

[0059] As mentioned hereinabove, the motion estimation module 32 allows for predicting an area of the previous frame that might have been moved into the current frame.

[0060] Then, the motion compensation module 34 aims at compensating for the movement detected by the motion estimation module 32.

[0061] The *DCT* module 36 performs the *DCT* transformation on blocks of pixels into spatial frequency coefficients.

[0062] Those spatial frequency coefficients are then supplied to the quantizing module 38 for quantizing them.

[0063] Encoded frames can be stored in a buffer 40 of the encoder 14. Therefore, the buffer 40 allows for receiving encoded frames, which are then transmitted through the communication link (network 16).

[0064] Finally, the encoder 14 also includes a rate controlling device 42 according to the present application, for controlling the bitrate in video coding.

This rate controlling device ensures that the maximum capacity of the communication link is not exceeded and provides enhanced visual quality, especially for tele-surveillance applications and other applications exhibiting continuous video sequences.

II. The device

[0065] The rate controlling device 42 implements an improved rate control in the encoder 14. More specifically, the device 42 implements a method 100 for controlling the bitrate in the *H.263* and *MPEG 4* video coding standard so as to meet with an average and maximum bitrate constraint. The method 100 will be described hereinbelow.

[0066] The device 42, as illustrated in Figure 3, includes a predictor 50 for specifying a target size for Intra frames and predicting the required QP used in the quantizing module 38, for meeting the specified Intra frame target size. Since the size of the Intra frames can affect both quality and bursts in the bitrate, when controlling the bitrate, a target size is determined to set properly a bit budget for encoding the Intra frames.

[0067] The device 42 also includes a first calculator 52 of Inter frame sizes. The first calculator 52 allows for computing a target size of the Inter frames in different situations, for meeting with a maximum bitrate constraint, as will be described hereinbelow.

[0068] Furthermore, the device 42 includes a second calculator 54 for recalculating the Inter frame target sizes provided by the first calculator 52, so as to meet both the maximum bitrate and an average bitrate constraint.

[0069] Also, the device 42 includes a validator 56 for ensuring that the size of the encoded frame does not exceed the calculated inter frame target size and the maximum size allowed by the maximum bitrate.

[0070] Finally, in order to meet with the constraint of the maximum and average bitrates, the process 30₂ in Figure 2 has a circular buffer (not shown) for containing the size of the frames transmitted within the last one (1) second. For instance, at 10 frames/second, such a buffer would keep the size of the last 10 encoded frames.

[0071] Generally stated, the method 100 according to a non-restrictive illustrative embodiment of the present invention, for meeting with the maximum and average bitrates constraint uses two control processes, such as 30₂ and 30₃. For example, the process 30₂ may be used, along with the first and second calculators 52 and 54, for computing the target frame sizes required to avoid exceeding the maximum bitrate. And the process 30₃, along with the validator 56, is used to check, after encoding a frame, if the encoded frame can be transmitted, i.e. the maximum bitrate is not reached. In case the maximum bitrate has been exceeded, the encoded frame is dropped or the encoder 14 attempts to encode again the same frame but with different encoding parameters.

III. The method

[0072] In the following, the method 100 according to the non-restrictive illustrative embodiment of the present invention for controlling the bitrate in video coding standards, such as the H.263 standard, will be described in connection with Figure 4.

[0073] However, it should be noted that this rate control can be also used in other video coding standards, such as *H.261*, *MPEG-4*, *H.264*, etc.

[0074] As mentioned hereinabove, since the size of the Intra frames can affect both the quality and bursts in the bitrates, it is important to set a first bit budget for encoding the Intra frames, when controlling the bitrates.

[0075] In operation 101, a maximum bitrate R_{MAX} and an average bitrate R_{avg} are specified.

[0076] Then, in operation 102, a target size B_I (in bits) of the Intra frame is specified.

[0077] In operation 103 of method 100, prediction of QP is performed, which allows for better meeting with the Intra frame target size specified in operation 102 when encoding an Intra frame. Operations 102 and 103 can be performed in the predictor 50 of Figure 3, for example.

[0078] Then, in operation 104, a target size for an Inter frame is calculated in order to meet the specified maximum bitrate R_{MAX} , based on the target size of the Intra frame specified in operation 102. This operation 104 can be performed by the first calculator 52 in Figure 3, for example.

[0079] In operation 106, the calculated Inter frame target size is recomputed so as to meet both the specified maximum bitrate and average bitrate. This operation can be performed by the second calculator 54 in Figure 3, for example.

[0080] Finally, in operation 108, encoded frame sizes are validated against the computed frame sizes allowed by the maximum and average bitrates.

[0081] Now, more specifically, each operation of the method 100 of Figure 4 will be described with greater details.

Operation 101: specification of a maximum bitrate and an average bitrate

[0082] In order to obtain an improved quality of videos in video coding, a maximum bitrate, denoted as R_{MAX} , whose value should not be exceeded is set. The maximum bitrate is defined as the maximum number of bits that can be transmitted through a channel within any period of one (1) second. With this definition, the maximum size is equivalent to the maximum bitrate. It should be noted that other definitions of the maximum rate can also apply. It is within the scope and knowledge of those of ordinary skill in the art to specify and define a maximum bitrate according to each application. It is also known in the art how to specify an average bitrate over a period of frames.

Operation 102: specification of an Intra frame target size

[0083] It is within the knowledge and reach of those of ordinary skill in the art to know how to specify an Intra frame target size based on a particular application of video coding, and on the specified maximum and average bitrates. For instance, the Intra frame target size could be set as a percentage of the maximum bitrate R_{MAX} or the average bitrate R_{avg} .

Operation 103: prediction of QP in order to meet the specified Intra frame target size

[0084] To do so, based on the estimation algorithm disclosed in “*Adaptive intra-frame quantization for very low bit rate video coding*” by Pan et al. 2004, hereinafter referred to as Reference 2, an estimation algorithm for predicting a QP (Quantization Parameter), used in the quantizing module 38, has been developed.

[0085] This algorithm is used by the predictor 50 in Figure 3 for estimating the QP required in the quantizing module 38, in order to meet with the given Intra frame target size, based on the statistics of the Intra frame. However, unlike the teaching of Reference 2, where a constant QP is considered for the whole duration of the Intra frame, in the estimation algorithm that has been developed, the QP value may be updated for each GOB (Group of Blocks) of the Intra frame so as to meet more accurately with the specified target frame size. Also, while Reference 2 considers gray scale images, the developed estimation algorithm supports color images.

[0086] According to Reference 2, the relationship between the QP and other frame statistics can be described by:

$$QP = \frac{16.34}{B_I^{2.05}} MAV_{DCT}^{1.0+0.29 \ln(B_I)} \quad (1)$$

with $MAV_{DCT} = \frac{1}{MN} \sum_{u=0}^{N-1} \sum_{v=0}^{M-1} ABS(F(u,v))$, $F(u,v)$ containing the 8×8 DCT coefficients of all blocks of pixels of the picture of size $N \times M$ and B_I is the specified Intra frame target size (in bits).

[0087] The term MAV_{DCT} represents the mean absolute value of the DCT coefficients, and indicates a picture complexity measure. The term $ABS()$ is the absolute value function.

[0088] It should be noted that equation (1) only considers the statistics of the luminance of a picture and ignores its chrominance. Therefore, a new equation is needed in order to take also into account the chrominance, which is generally subsampled by a factor of four (4) compared to the luminance, and is given by:

$$MAV_{DCT} = \frac{0.9}{MN} \sum_{u=0}^{N-1} \sum_{v=0}^{M-1} ABS(F_L(u, v)) + \frac{0.1}{4MN} \left(\sum_{u=0}^{N/2-1} \sum_{v=0}^{M/2-1} (ABS(F_{C1}(u, v)) + ABS(F_{C2}(u, v))) \right) \quad (2)$$

where $F_L(u, v)$ contains the 8×8 DCT coefficients of the blocks for the luminance and $F_{C1}(u, v)$ and $F_{C2}(u, v)$ contain the 8×8 DCT coefficients of the blocks for the chrominance.

[0089] As mentioned hereinabove, in order to meet more accurately with the specified Intra frame target size B_I , the QP is adjusted for each GOB of the Intra frame. Furthermore, equation (2) is modified so as to take into account the bits needed for encoding the remaining blocks of the Intra frame. The modified equation for predicting the QP is then as follows:

$$QP_{\text{mod}} = \frac{16.34}{B_I^{2.05}} MAV_{DCT}^{1.0+0.29 \ln(B_I)} \quad (3)$$

with $B_I = \frac{(B_{\text{remain}})N_{\text{total}}}{(N_{\text{remain}} + N_{\text{bias}})}$, where B_{remain} is the number of bits that remain from the specified Intra frame target size for encoding the frame, N_{total} is the total number of macroblocks in a frame, N_{remain} is the number of the remaining macroblocks to encode in the frame, and N_{bias} is a bias for making the QP estimation more conservative as the end of the frame to encode is approaching; typically $N_{\text{bias}} = 4$.

[0090] After the QP have been computed so as to meet with the specified Intra frame target size B_I , method 100 moves on to operation 104.

Operation 104: calculation of an Inter frame size in order to meet the maximum bitrate constraint

[0091] As a non-restrictive example, the Intra frame target size specified in operation 101 is considered constant; however the Inter frame target size depends on the situation. Two cases need to be considered:

- i) Case 1: when an Intra frame has been encoded within the last second or will be encoded within the upcoming second; and
- ii) Case 2: only Inter frames have been encoded within the last second or will be encoded in the upcoming second, without any influences from the Intra frames.

[0092] Therefore, operation 104 needs to consider the two (2) above-mentioned cases, namely case 1 and case 2, for computing the Inter frame target size. Method 200 is used in case 1 and method 250 is used in case 2 by operation 104.

Method 200 for calculating the Inter frame target size in the case where an Intra frame is encoded within the last second or is going to be encoded in the upcoming second (case 1)

[0093] Referring to Figure 5, method 200 for calculating the Inter frame target size, in case 1, will be described.

[0094] First, in operation 202, an ideal Inter frame target size B_2 is computed, based on R_{MAX} . To do so, as a non-limitative example, all the Inter frames are assumed to have the same size, then, the maximum bitrate is given by:

$$R_{MAX} = B_I + (F - 1) B_2 \quad (4)$$

with B_I being the Intra frame target size and B_2 being the ideal Inter frame target size in case 1, i.e. with an Intra frame encoded within the last second or to be encoded in the upcoming second and F being the number of frames per second to encode, i.e. the frame rate.

[0095] Then, from equation (4) the ideal Inter frame target size B_2 is given by:

$$B_2 = \frac{R_{MAX} - B_I}{(F - 1)} \quad (5)$$

By definition, as shown in equation (5), F is greater than 1 frame/s.

[0096] At first, the Inter frame target size for case 1, denoted as B_{PI} , is given by the ideal Inter frame target size B_2 calculated in operation 202:

$$B_{PI} = B_2$$

[0097] However, the above Inter frame target size B_{PI} is adjusted to take into account the influences of the past encoded frames, especially the Intra frames. Indeed, in general, the encoded frames do not have a size which is exactly equal to the target size. Accordingly, the future Inter frame sizes are adjusted so that the differences between the actual size of the encoded frames and the target sizes are accounted for so as to ensure that there will be enough bits allocated in the encoding of a future Intra frame, without exceeding the maximum bitrate R_{MAX} .

[0098] Therefore, in operation 204, an adjusted Inter frame target size based on the past encoded frames is computed, through the first calculator 52 of Figure 3, for example. To do this, two (2) cases are considered:

[0099] i) If an Intra frame has been encoded within the last second, the size of the next Inter frame B_{PI} will be adjusted and limited by using the following equation:

$$B_{PI} = \max(0, \min(B_{PI}, tmp)) \quad (6)$$

with

$$tmp(t) = \frac{R_{MAX} - \sum_{k=t-(T-N)}^{t-1} SizeHistory[k]}{F - (T - N)} \quad (7)$$

where $SizeHistory[k]$ is the size of the encoded frame k , t is the index of the current frame to encode, T is the period (in number of frames) between two Intra frames, N is the number of frames remaining before the next Intra frame. $tmp(t)$ represents the bit budget buffer at frame t . Also, in equation (6), the zero (0) inside the maximum function allows for avoiding an underflow of the buffer.

[00100] It should be noted that equations (6) and (7) allow for dividing substantially equally any additional bits from the bit budget for encoding frames between the remaining Inter frames to be encoded within the second following the Intra frame. Also, it should be noted that the frame at time $t-(T-N)$ is the one which was Intra-coded.

[00101] ii) If an Intra frame is to be encoded in the upcoming second then the target size of the next Inter frame will be limited as follows:

$$B_{pI} = \max(0, \min(B_{pI}, tmp)) \quad (8)$$

with

$$tmp(t) = \frac{R_{MAX} - B_I - \sum_{k=t-(F-N)+1}^{t-1} SizeHistory[k]}{N} \quad (9)$$

where $SizeHistory[k]$ is the size of the encoded frame k , t is the index of the current frame to be encoded, and N is the number of frames remaining before the next Intra frame.

[00102] It should be noted that in equation (9), $F-(N-1)$ terms have been added. Therefore, the size of the $(N-1)$ remaining frames is calculated in the period with one frame being an Intra frame. Accordingly, the equation representing the bit budget buffer $tmp(t)$ contains a division by N .

[00103] Finally, in operation 206, the adjusted Inter frame target size is recalculated so as to ensure that it does not exceed the maximum size allowed by the maximum bitrate R_{MAX} , while keeping a small margin of error and considering the bits that have been used to encode the past frames.

[00104] Accordingly, the Inter frame target size is limited by using the following equation:

$$B_{PI} = \max(0, \min(B_{PI}, tmp2)) \quad (10)$$

$$\text{with } tmp2(t) = R_{MAX} - R_{margin} - B_I - \sum_{k=F+1}^{t-1} \text{SizeHistory}[k] \quad (11)$$

where R_{margin} is the margin of error for ensuring that the Inter frame target size does not exceed the maximum size allowed by the maximum bitrate R_{MAX} .

[00105] In case there is no Intra frame encoded within the last second or to be encoded in the upcoming second, which corresponds to case 2, operation 104 uses method 250 for calculating an Inter frame target size.

Method 250 for calculating the Inter frame target size in the case where no intra frame is encoded within the last second or to be encoded in the coming second (case 2)

[00106] Turning now to Figure 6, method 250 for calculating an Inter frame target size, in case 2, will be described.

[00107] First, in operation 252, an ideal target size B_3 of the Inter frames is calculated, based on the maximum bitrate. In this case, since it is assumed that the Intra frames are absent, the maximum bitrate R_{MAX} during any second is given by:

$$R_{MAX} = F \cdot B_3 \quad (12)$$

where B_3 is the ideal Inter frame target size without any Intra frame encoded within the last second or to be encoded in the upcoming second.

[00108] Therefore, the ideal Inter frame target size B_3 is given by:

$$B_3 = \frac{R_{MAX}}{F} \quad (13)$$

[00109] At first, the Inter frame target size, denoted as B_P , in case 2, takes the value of the calculated ideal Inter frame size B_3 , i.e. $B_P = B_3$.

[00110] In this case, there is no need of adjusting the target size B_P based on the past encoded Intra frames, since it is assumed that the Inter frames are not influenced by the Intra frames.

[00111] Then, in operation 254, the Inter frame target size B_P is recalculated so as to ensure that it does not exceed the maximum size allowed by the maximum bitrate R_{MAX} , while keeping a small margin of error and taking into consideration the bits used for encoding the past frames.

[00112] To do this, the following expression is used:

$$B_P = \max(0, \min(B_P, \text{tmp2})) \quad (14)$$

with $\text{tmp2}(t) = R_{MAX} - R_{margin} - \sum_{k=t-F}^{t-1} \text{SizeHistory}[k]$, where R_{margin} is the margin of error between the target size and maximum size allowed by the maximum bitrate; for example, $R_{margin} = 1000$ bits.

[00113] Once the Inter frame target size has been calculated, either in case 1 or case 2, for meeting with the maximum bitrate constraint, method 100 in Figure 4 moves on to operation 106, for updating and adjusting the calculated Inter frame size so as to meet both the maximum bitrate and the average bitrate specified in operation 101. This operation 106 can be performed in the second calculator 54 of Figure 3, for example.

Operation 106: re-calculation of the calculated Inter frame size in order to meet both an average and maximum bitrate constraint

[00114] Operation 106 uses methods which allow for meeting with both average and maximum bitrate constraint. These methods have two components: i) the computation of Inter frame target sizes in order to meet with the average and maximum bitrate constraint, and ii) the management of additional bits that have not been used in the encoding of the past frames and which could be used in the encoding of the future frames.

[00115] Again, the two cases, case 1 and case 2 as described in operation 104, will be considered. Therefore, operation 106 uses method 300 in case 1 and method 350 in case 2.

Method 300 for re-calculating the Inter frame size in the case where an intra frame is encoded within the last second or is going to be encoded in the upcoming second (case 1)

[00116] Referring now to Figure 7, method 300 for re-calculating the Inter frame target size calculated in operation 104, in case 1, will be described.

[00117] First, in operation 302, an ideal Inter frame target size B is calculated, based on the specified average bitrate.

[00118] Ideally, Inter frames have the same size, so that, over a period T between two Intra frames, the average bitrate R_{avg} is given by:

$$R_{avg} = \frac{F(B_I + (T - 1)B)}{T} \quad (15)$$

[00119] Thus, the ideal Inter frame target size B can be expressed as follows:

$$B = \frac{(R_{avg} T / F) - B_I}{T - 1} \quad (16)$$

[00120] In order to ensure that the Inter frame size does not exceed the average bitrate R_{avg} and the maximum bitrate R_{MAX} , the Inter frame target size B_{PI} is recalculated as follows:

$$B_{PI} = \min(B, B_2) = \min\left(B, \frac{R_{MAX} - B_I}{F - 1}\right) \quad (17)$$

[00121] Operation 302 can be computed only once, either when the encoder 14 is first initialized or each time that the bitrate parameters change; the parameters include the specified maximum bitrate, for example.

[00122] Then, in operation 304, the calculated target frame size B_{PI} is adjusted so as to take into account the past encoded frames, more particularly the Intra frames. To do so, equations similar to equation (6) or (8) are

calculated for obtaining the updated Inter frame target size B_{PI} , based on the past encoded frames:

$$B_{PI} = \max(0, \min(B_{PI}, tmp))$$

with tmp as defined in equation (7) or (9) depending on the case: i) if an Intra frame has been encoded within the last second, then equation (7) is used and ii) if an Intra frame is to be encoded in the upcoming second, then equation (9) is used.

[00123] Next, in operation 306, differences between the adjusted target frame size and the actual frame size are calculated since, as mentioned hereinabove, encoded frames do not generally have exactly the same size as the target size.

[00124] To do so, a variable B_{extra} is used to monitor these differences through the use of the process 30₁ of Figure 2, for example. This variable is initialized to zero at the beginning of the communication session. The variable B_{extra} accumulates the number of bits which are over or under the target size, from frame to frame. If B_{extra} is greater than zero, then the actual bitrate is under the specified average bitrate; if B_{extra} is negative, then the actual bitrate is over the specified average bitrate. The update of B_{extra} is performed for each encoded frame as follows:

$$B_{extra} = B_{extra} + B_{target} - B_{encoded} \quad (18)$$

where B_{target} is the target size, which corresponds to B_{PI} for an Inter frame and B_I for an Intra frame, and $B_{encoded}$ is the size of the encoded frame.

[00125] If B_{extra} is positive, meaning that the actual bitrate is under the specified average bitrate, then, the unused bits are distributed over the next frames to be used in their encoding; therefore, the target size is increased for the next frames. In the case where B_{extra} is negative, the target size is reduced, for the next frames. The algorithm used is as follows:

If $B_{extra} > 0$ (this corresponds to the case where the actual bitrate is under the average bitrate), then

$$B_{target} = B_{target} + \alpha B_{extra} \quad \text{if} \quad \alpha B_{extra} \leq \beta B_{target}$$

or

$$B_{target} = B_{target} + \beta B_{target} \quad \text{if} \quad \alpha B_{extra} > \beta B_{target}$$

Otherwise

$$B_{target} = B_{target} + B_{extra}$$

with α representing the fraction of additional bits that are used in the current frame and β representing the largest fraction of the target size that can be added. For example, in the simulations, $\alpha=0.6$ and $\beta=0.15$ are used. However, these values could be modified to further improve the performance of the encoder 14. It is within the scope and knowledge of a person of ordinary skill in

the art to adjust the values of α and β in order to obtain various performances. Furthermore, the above described method for updating the target size B_{target} can be also modified or changed without deviating from the main concept of this rate control method.

[00126] It should be noted that, from an implementation perspective, it is possible to drop encoded frames which lead to a negative B_{extra} . In such a case, the actual average bitrate of the encoded frames would be over the specified average bitrate. Also, it is possible to decide to drop encoded frames, which are below a certain threshold. For example, an implementation can use a threshold value of -500 bits.

[00127] Finally, operation 308 makes sure that the maximum size allowed by the maximum bitrate R_{MAX} is not exceeded by the adjusted Inter frame target size for the current frame to be encoded. To do so, an equation similar to equation (10) is used to obtain a valid Inter frame target size:

$$B_{pI} = \max(0, \min(B_{pI}, tmp2))$$

with $tmp2$ as defined in equation (11).

[00128] It is also possible to modify method 300 so as to account for B_{extra} right before operation 306. This is an implementation choice which would lead to some other compromises in behavior and performance without deviating from the main ideas of this rate control method.

[00129] In case there is no Intra frame that has been encoded within the last second or that is going to be encoded in the upcoming second, operation 106 uses method 350 for updating the calculated Inter frame target size in order to meet with both the average and maximum bitrates.

Method 350 for re-calculating the Inter frame size in the case where no Intra frame is encoded within the last second or is going to be encoded in the upcoming second (case 2)

[00130] As mentioned hereinabove, only ideally the Inter frames are of the same size. In reality, the maximum bitrate constraint may have reduced significantly the target size of the Inter frames located close to the Intra frames. In order to compensate for these differences in size, the size of the Inter frames that are not influenced by the Intra frames will be re-evaluated.

[00131] Referring now to Figure 8, method 350 for re-calculating the Inter frame target size will be described, in case 2.

[00132] First, in operation 352, an ideal Inter frame target size B_a is calculated, based on the average bitrate.

[00133] In the ideal situation, the size of the Inter frames is assumed to be of the same length. Therefore, in that case, over the whole period T between two Intra frames, the average bitrate R_{avg} is given by:

$$R_{avg} = F \left(\frac{B_I + (2F - 2)B_{Pl} + (T - (2F - 1))B_a}{T} \right) \quad (19)$$

[00134] From equation (19), it can be deduced that the ideal size B_a of the Inter frames is given by:

$$B_a = \frac{(R_{avg} T / F) - B_I - (2F - 2)B_{pI}}{T - (2F - 1)} \quad (20)$$

[00135] Operation 352 can be computed only once, either when the encoder 14 is initialized or each time the bitrate parameters change.

[00136] In order to ensure that the Inter frame target size B_P does not exceed the ideal Inter frame target size based on the average bitrate and the Inter frame size calculated in operation 252 of method 250 for not exceeding the maximum bitrate, the target size B_P of the Inter frame is limited as follows:

$$B_P = \min(B_a, B_3) = \min\left(B_a, \frac{R_{MAX}}{F}\right) \quad (21)$$

[00137] Then, in operation 354, a difference between the Inter frame size B_P computed in operation 352 and the actual encoded frame size is calculated. This operation is performed for each frame.

[00138] The number of bits, resulting from the difference computed in operation 354, are tracked and stored in the variable B_{extra} . This variable allows for determining if the actual encoded frame is over or under the specified average bitrate and for compensating for the differences based on the situation. This is done as described in operation 306 of method 300 of Figure 7 with the target size calculated according to equation (21).

[00139] Finally, in operation 356, B_P is limited so as to not exceed the maximum bitrate, while keeping a margin of error and considering the bits used for encoding the past frames. B_P is given by the following equation:

$$B_P = \max(0, \min(B_{pl}, \text{tmp2})) \quad (22)$$

with $\text{tmp2} = R_{MAX} - R_{margin} - \sum_{k=t-F}^{t-1} \text{SizeHistory}[k]$, and R_{margin} is the margin of error for the bitrate. For example, $R_{margin} = 1000$ bits.

[00140] Once the Inter frame target sizes are calculated, in operation 108 of Figure 4, a validation operation is performed by the validator 56 of Figure 3, for example. Indeed, encoded frames usually do not have the same size as the calculated Inter frames. Therefore, if the encoded frame sizes are too large when compared to the calculated target sizes and maximum bitrate, then a re-encoding or dropping mechanism is activated. Those mechanisms are well known in the art and therefore will not be further described in the present specification.

IV. Numerical examples

[00141] In this section, some numerical examples will be described according to method 100 for rate controlling in video coding.

[00142] For example, according to method 200 and method 250, if the maximum bitrate is $R_{MAX} = 48000$ bps (bits per second), the Intra frame target size is $B_I = 40000$ bits and the frame rate is $F = 10$ fps (frames per second) then:

By using equation (5) in operation 202 and equation (12) in operation 252, respectively, it can be found that:

$$B_I = 40000$$

$$B_{PI} = \frac{48000 - 40000}{9} = 888 \text{ bits}$$

$$B_P = \frac{48000}{10} = 4800 \text{ bits}$$

[00143] Again, operations 202 and 252 can be computed only once, either when the encoder 14 is initialized or each time that the bitrate parameters change. However, operations 204, 206 and 254 are computed each time, i.e. for each frame, based on the actual encoded frame sizes.

[00144] Another example is concerned with method 300 and method 350. In this case, if the maximum bitrate is $R_{MAX} = 48000 \text{ bps}$, the average bitrate is $R_{avg} = 32000 \text{ bps}$, the Intra frame target size is $B_I = 40000 \text{ bits}$, the period of the Intra frames is $T = 4s$, i.e. each 40 frames, and the frame rate is $F = 10 \text{ fps}$, then:

By using equation (17) in operation 302 and equation (20) in operation 354, it can be found that:

$$B_I = 40000$$

$$B_{PI} = \min \left(\frac{\left(\frac{32000.40}{10} \right) - 4000}{39}, \frac{48000 - 40000}{9} \right) = \min(2256, 888) = 888 \text{ bits}$$

$$B_P = \min \left(\frac{\left(\frac{32000.40}{10} \right) - 40000 - (20 - 2)(888)}{40 - 19}, \frac{48000}{10} \right) = \min(3429, 4800) = 3429 \text{ bits}$$

[00145] Again, operations 302 and 352 can be computed only once, either when the encoder 14 is initialized or each time that the bitrate parameters change. However, operations 304, 306, 308, 354 and 356 are computed for each frame, based on the actual encoded frame sizes.

[00146] The average bitrate, which can be obtained if all frames are encoded as planned by using the calculated target sizes, is then computed, using equation (19) with $B_P = 3429$ bits:

$$\begin{aligned} R_{avg} &= F \left(\frac{B_I + (2F - 2)B_{PI} + (T - (2F - 1))B_P}{T} \right) \\ &= 10 \left(\frac{40000 + (18)(888) + (40 - (19))(3429)}{40} \right) \\ &= 32000 \text{ bits} \end{aligned}$$

[00147] Therefore, the desired ideal result is obtained. In certain cases, the maximum bitrate and Intra frame target size may limit the average bitrate especially if the maximum bitrate is close to the average bitrate and if

the target Intra frame size is large. For instance, if in the previous example the average bitrate is *48000 bps* then:

$$B_I = 40000$$

$$B_{PI} = \min \left(\frac{\left(\frac{48000 \cdot 40}{10} \right) - 4000}{39}, \frac{48000 - 40000}{9} \right) = \min(3897, 888) = 888 \text{ bits}$$

$$B_P = \min \left(\frac{\left(\frac{48000 \cdot 40}{10} \right) - 40000 - (20 - 2)(888)}{40 - 19}, \frac{48000}{10} \right) = \min(6476, 4800) = 4800 \text{ bits}$$

[00148] The average bitrate, which can be obtained if all frames are encoded as planned, is computed as follows, with $B_P = 4800 \text{ bits}$:

$$\begin{aligned} R_{avg} &= F \left(\frac{B_I + (2F - 2)B_{PI} + (T - (2F - 1))B_P}{T} \right) \\ &= 10 \left(\frac{40000 + (18)(888) + (40 - (19))(4800)}{40} \right) \\ &= 39196 \text{ bits} \end{aligned}$$

[00149] But in reality, a smaller average bitrate is observed because encoded frame sizes will fluctuate around the target sizes. Therefore, results given by equation (19) correspond to a theoretical maximum.

[00150] It should be noted that the calculation of the Inter frame target size in method 100 depends on the position of the Inter frame relative to an Intra frame.

V. Management of encoded frames leading to exceeding the calculated target sizes

[00151] In the TMN8 rate control, all encoded frames are transmitted even if their transmission will lead to a situation where the specified average bitrate is exceeded. After the transmission of a frame, if the specified average bitrate is exceeded, then the encoder will skip one or several frames, as needed, until the specified average bitrate is respected again. Therefore this control is reactive rather than proactive.

[00152] The above scheme may be acceptable for meeting an average bitrate. However, it is not acceptable if a maximum bitrate must never be exceeded, as it is the case in many real life applications using a physical transmission channel.

[00153] Indeed, since the size of any frame can potentially exceed the target size, a mechanism where encoded frames leading to exceeding the maximum bitrate will be dropped before being transmitted should be implemented. The problem with this approach is that video encoding uses frame to frame prediction (temporal predictive coding). Therefore, if a frame is dropped, the state of the encoder needs to be restored prior to the re-encoding of the dropped frame, otherwise the encoder and decoder will not be synchronized.

[00154] Therefore, the device 42 according to the non-restrictive embodiment of the present invention memorizes the encoder state variables before encoding a frame. If the encoded frame size leads to exceeding the maximum bitrate, then it is dropped and the state variables are restored for the next frame to encode. Otherwise the encoded frame is transmitted and the encoder 14 continues with the new state variables.

[00155] This feature can also be used for other applications such as multipass encoding where multiple encoding iterations on a frame can be performed until a goal in size and/or quality are met.

[00156] The encoder 14 may drop or skip frames under two conditions. First, if an encoded frame exceeds the maximum bitrate then it will be automatically dropped. Secondly, if the number of bits transmitted exceeds the specified average bitrate, then future frames may be skipped, not the one that was just encoded but the ones that follow, until the average bitrate is respected again. More specifically, future frames could be skipped if the specified average bitrate is exceeded by a certain threshold value and the current frame will be dropped only in the case where keeping it would lead to exceeding the maximum bitrate.

[00157] Furthermore, the encoder 14 allows for requesting an Intra frame at any moment, leading to an implementation of an “Intra frame on demand” feature. Obviously, in such a case, the Intra frame target size may not be respected without exceeding the maximum bitrate. If this happens, the encoder 14 will compute and use the maximum Intra frame target size without exceeding the maximum bitrate. If the encoded Intra frame size exceeds the maximum bitrate, then it will be dropped and the encoder 14 will wait for a certain period of time before trying again to encode that Intra frame. The encoder 14 will continue to wait and try to encode that Intra frame until the encoded Intra frame size respects the maximum bitrate. At each failed attempt, the Intra frame target size is reduced so as to avoid an infinite loop, where the encoded frame size is always above the maximum bitrate.

[00158] In general, the encoder 14 requests the following parameters, when encoding video data and frames:

- Intra frame target size, as specified in operation 102 of method 100;

- Average bitrate, as specified in operation 101 of method 100; and
- Maximum bitrate, as specified in operation 101 of method 100.

[00159] It should be noted that the encoder 14 can perform several validations of the above-given parameters and may change the parameter values if using them is not possible. For instance, with a given Intra frame target size and maximum bitrate, the average bitrate may not be achievable. Therefore the parameters will be changed and the average rate recalculated.

VI. A second non-limitative embodiment of the rate control for H.263 based on a buffer

[00160] In the previous embodiment, the maximum bitrate R_{MAX} was defined as the maximum number of bits which can be transmitted during a period of one (1) second.

[00161] However, in applications such as video streaming, a different definition of the maximum bitrate is considered, which takes into account the actual limitations existing in those applications.

[00162] In video streaming, video packets are sent through a capacity-limited channel; therefore the bitrate is limited. If a delay of one (1) second is considered acceptable before the player 20 in Figure 1 starts playing a first received Intra frame, then the Intra frame target size should be inferior to R_{MAX} as will be explained hereinbelow.

[00163] The decoder 18 has to wait for one (1) second before the player 20 starts playing the first received frame. However, it does not mean that one cannot transmit additional frames during this delay of one (1) second. Now, suppose that $B_t = R_{MAX}$ and that not exceeding the maximum bitrate R_{MAX} is

equivalent to the condition that the transmission buffer 40 does not exceed R_{MAX} . Then, at time $t = 0$, the buffer 40 of Figure 2 is full, since a frame of size R_{MAX} has been just inserted. If $F = 10$ fps and $R_{MAX} = 48000$ bps, then at $t = 0.1$ second, the buffer 40 is emptied by $R_{MAX}/F = 4800$ bits, since they have been transmitted through the channel. Therefore, 4800 bits become available for encoding a frame at the given time of $t = 0.1$ second and without overflowing the transmission buffer 40. The same approach applies for the following frames. Therefore, for each frame following the Intra frame, an additional amount of 4800 bits can be transmitted. Thus, with this approach, it is possible to transmit in total $48000 + (10-1)(4800) = 91200$ bits, instead of 48000 bits as shown and described hereinabove.

[00164] Those additional available bits allow for improving the quality of the frames following an Intra frame, since they are used to increase the number of bits for encoding the Inter frames following the Intra frame.

[00165] However, it should be noted that the 91200 bits calculated previously, are not sent during one (1) second but because of the delay induced by the applications to receive a transmitted Intra frame, the time covering those bits is 1.9 seconds: 1 second of delay and 9 frames from time 1.1s to 1.9s, which gives a bitrate of $(91200 \text{ bits} / 1.9\text{s}) = 48000 \text{ bits/s}$, corresponding to the maximum bitrate R_{MAX} .

[00166] For illustration purposes, it will be considered that the maximum buffer size will correspond to the maximum bitrate R_{MAX} . However the maximum buffer size could be set with a different value. Those of ordinary skill in the art will know how to modify the formulas accordingly.

[00167] Now, the case where a next Intra frame is coming will be considered. Since the next Intra frame is coming, the buffer 40 has to be

gradually emptied in order to have sufficient space for the upcoming Intra frame at its scheduled time.

[00168] Even though the target size for the Intra frame is B_I , it is desirable, for practical purposes, to have an empty buffer 40 so as to allocate more available bits to the frames following the Intra frame in order to improve quality thereof.

[00169] So, in practical applications, $B_I < R_{MAX}$. Therefore, the number of bits allocated to the frames following the Intra frame is increased so as to obtain an improved quality of those frames, before starting to empty the buffer 40.

[00170] Let zero (0) be the time of the arrival of an Intra frame. Then, N is defined as the number of frames after the Intra frame over which the unused bits from the encoding of the Intra frame are distributed. And L is defined as the number of Inter frames over which the buffer level is optimized. It should be noted that at $t = L$, the buffer 40 is full and $0 \leq N \leq L$.

[00171] Generally stated, a possible strategy that can be adopted is as follows.

[00172] At all times, the buffer level does not exceed R_{MAX} , meaning that encoded frames that generate an overflow will be dropped. The bits not used in the encoding of the Intra frame are used in the encoding of the following Inter frames in order to maximize the buffer level and to increase the overall quality of the videos, since usually the quality of videos is the worst after an Intra frame. Until frame number L , the buffer 40 is kept full by encoding frames with a target size of R_{MAX}/F . Furthermore, the bits that have been

computed but not used in the previous frames are redistributed over the following frames; however the target size must not exceed a certain value, which depends on the target buffer level, which itself depends on the frame index in time. It should be noted that the target buffer level is R_{MAX} for the first frame after an Intra frame and gradually decreases to zero (0) until the next Intra frame comes.

[00173] Referring now to Figure 9, method 400 for improved rate controlling in video coding will be described.

[00174] In operation 402, at time $t = 0$, an Intra frame is received by the transmission buffer 40. Therefore, at time $t = 0$, the target frame size is $B_P(t) = B_I$.

[00175] In operation 404, comparison between a new buffer level (which is the sum of the previous buffer level and the size of the actual encoded Intra frame) and the maximum buffer size R_{MAX} is performed.

[00176] If the new buffer level is larger than the maximum buffer size R_{MAX} as tested in operation 406, then in operation 408, the actual encoded frame is dropped. Also the encoder state variables prior to encoding that frame will be restored (including the buffer level, which will be restored to the buffer level value prior to encoding).

[00177] If the new buffer level is smaller than the maximum buffer size R_{MAX} , then in operation 410, an Inter frame target size $B_p(t)$ is calculated for each frame at time t within a period T between two Intra frames, by using the following expression:

$$B_p(t) = B_p^I(t) \quad \text{if } t = 1, 2, 3, \dots, L$$

$$B_p(t) = \frac{TR_{MAX} - FB_I - F \sum_{k=1}^L B_p^I(k)}{F(T - L - 1)} \quad \text{if } t = L, L+1, \dots, T-1$$

with $B_p^I(t)$ being arbitrary but with the condition that the buffer 40 must not overflow, i.e. the buffer level $\leq R_{MAX}$ and has to be near full at $t = L$.

[00178] Then, in operation 412, a maximum buffer level is calculated, which cannot be exceeded. Indeed, for each frame, the maximum buffer level should be computed so as to ensure that there is sufficient space in the buffer 40 for inserting the next Intra frame. The size of an Intra frame cannot exceed R_{MAX} . Therefore, for the frame at $T-1$, the buffer level cannot exceed $R_{MAX} - B_I + R_{MAX}/F$ so that at the beginning of the next Intra frame, the buffer level does not exceed $R_{MAX} - B_I$.

[00179] However, for practical reasons, the size of the buffer 40 should be R_{MAX}/F instead of $R_{MAX} - B_I + R_{MAX}/F$, as mentioned hereinabove. This is because, in reality, it is hard to predict precisely the actual size of the encoded Intra frame. The current existing algorithms are not robust enough. Therefore, it is safer to have the buffer 40 be emptied at frame $T-1$ so that more bits can be used for the frames following the Intra frame.

[00180] Therefore, the maximum buffer level is calculated as follows:

$$BufferLevel_{MAX}(t) = R_{MAX} \quad \text{if } t = 0, 1, \dots, L$$

$$BufferLevel_{MAX}(t) = \frac{R_{MAX}(FT - F - L)}{F(T - L - 1)} + t \frac{R_{MAX}(1 - F)}{F(T - F)} \quad \text{if } t = L+1, \dots, T-1$$

[00181] Then, in operation 414, the Inter frame target size, calculated in operation 410, is updated so as to incorporate the unused bits from the encoding of the Intra frame.

[00182] There are many ways to distribute the unused bits over the Inter frames following an Intra frame. It should be noted that after frame L , the buffer 40 should be full in order to optimize the quality of the videos following the Intra frame.

[00183] For example, in order to distribute equally the unused bits over the N following frames, the target frame size at time t is given by:

$$B_p(t) = \frac{R_{MAX}}{F} + \frac{(R_{MAX} - B_I)}{N} \quad \text{if } 0 < t \leq N$$

$$B_p(t) = B_p^I = \frac{R_{MAX}}{F} \quad \text{if } N < t \leq L$$

$$B_p(t) = B_p^P = \frac{R_{MAX}(T - F - L)}{F(T - L - 1)} \quad \text{if } t = L, L+1, \dots, T-1$$

[00184] It should be noted that if $L = N$, then the target frame sizes of R_{MAX}/F are not present. Also, $T > F+L$ so that the buffer 40 will have sufficient time to empty itself between two Intra frames.

[00185] In operation 416, the buffer level, assigned to each future frame, is calculated so as to take into consideration the updated Inter frame target size.

[00186] Accordingly, the buffer level is given by:

$$BufferLevel(t) = R_{MAX} - \frac{(N - t)(R_{MAX} - B_I)}{N} \quad \text{if } 0 \leq t \leq N$$

$$BufferLevel(t) = R_{MAX} \quad \text{if } N \leq t < L$$

$$BufferLevel(t) = \frac{R_{MAX}(FT - F - L)}{F(T - L - 1)} + t \frac{R_{MAX}(1 - F)}{F(T - F)} \quad \text{if } t = L + 1, \dots, T - 1$$

[00187] It should be noted that, at the beginning, the transmission buffer 40 is empty and then frames of size B_I and $B_P(t)$ are successively inserted according to their computation.

[00188] Then, in operation 418, a margin $\Delta(t)$ for encoding a frame is calculated:

$$\Delta(t) = BufferLevel_{MAX}(t) - BufferLevel(t)$$

so that:

$$\Delta(t) = \frac{(N - t)(R_{MAX} - B_I)}{N} \quad \text{if } 0 \leq t \leq N$$

$$\Delta(t) = 0 \quad \text{if } N \leq t < T-1$$

[00189] Generally, when the encoder 14 encodes the frame, the encoded frame size is slightly different than the target size; in case a target size is larger than the encoded frame, the unused bits are redistributed over the next frames, as long as the maximum buffer level is not exceeded. Therefore, after the encoding of the frame, $E(t)$, indicative of the number of possible additional bits is computed in operation 420. Those unused bits, resulting from the encoding of the frame, can be used for encoding the next frames. $E(t)$ is defined as:

$$E(t) = \text{BufferLevel}(t) - \text{BufferLevel}_{\text{actual}}(t) = \\ \text{BufferLevel}_{\text{MAX}}(t) - \text{BufferLevel}_{\text{actual}}(t) - \Delta(t)$$

where $\text{BufferLevel}_{\text{actual}}(t)$ is the buffer level after the actual encoding of a frame.

[00190] It should be noted that generally a fraction of the additional bits is used to ensure that the maximum buffer level is not exceeded. Those of ordinary skill in the art will know how to specify the fraction of the additional bits in order to ensure that the maximum buffer level is not exceeded.

[00191] However, it may happen that an encoded frame is larger than a target frame size. In this case, $E(t)$ is negative or is inferior to a certain given threshold value. Therefore, in operation 422, a validation process is performed on the encoded frame. If the encoded frame has exceeded the target frame size, then this encoded frame is dropped. Also, in this case, the encoder state variables are restored to their values prior to encoding the frame.

[00192] It is worth noting that, for example, operation 410 of method 400 can be performed through the first calculator 52 in Figure 3, and operation 412 of method 400 can be performed through the second calculator 54 in Figure 3. Also, an optimizer (not shown) provided by the process 30₁ in Figure 2, for example, allows for optimizing the transmission buffer 40 over L frames.

VII. Numerical examples

[00193] For example, the same values as given in the previous numerical example can be used: $R_{MAX} = 48000$ bits, $B_i = 40\ 000$ bits, $F = 10$ fps, $T = 40$, i.e. 4 s. Furthermore, assuming that $N = 3$ and $L = F - 1 = 9$, then,

$$B_P(t) = 40\ 000 \quad \text{if } t = 0$$

$$B_P(t) = 7466 \quad \text{if } 0 < t \leq N$$

$$B_P(t) = 4800 \quad \text{if } N < t \leq L$$

$$B_P(t) = 3360 \quad \text{if } t = L, L + 1, \dots, T - 1$$

[00194] Using the above-calculated values, the maximum bitrate can be calculated as follows :

$$R_{MAX} = 10 \left(\frac{40000 + (3)(7466) + (9 - 3)(4800) + (39 - 9)(3360)}{40} \right) = 48000 \text{ bits}$$

[00195] This result is more interesting than the 39.2 kbps given by method 100. Also, since the additional bits are distributed over the Inter frames following an Intra frame, the visual quality of the frames is improved.

[00196] Although the present invention has been described in the foregoing specification by means of non-restrictive illustrative embodiments, these illustrative embodiments can be modified at will within the scope of the appended claims.

What is claimed is:

1. A method for rate controlling in video coding of sequences of pictures, including a series of Inter frames separated by Intra frames, the method comprising:
 - (a) for an Inter frame in the series, determining a maximum buffer level of a transmission buffer as a function of a position of the Inter frame relative to a previous Intra frame and an upcoming Intra frame in the sequence of pictures;
 - (b) determining a target frame size measured in bits for the Inter frame leading to a transmission buffer level not exceeding the maximum buffer level;
 - (c) encoding the Inter frame into an encoded frame according to the target frame size, the encoded frame having an encoded frame size;
 - (d) determining the transmission buffer level in response to the encoded frame; and
 - (e) dropping the encoded frame if it would cause the maximum buffer level to be exceeded.
2. The method of claim 1, the step (e) further including restoring the transmission buffer level to a previous value, provided the encoded frame was dropped.
3. The method as defined in any one of claims 1 or 2, wherein the step (a) comprises determining the maximum buffer level, which ensures that a transmission buffer can accommodate the upcoming Intra frame.
4. The method as defined in any one of claims 1 to 3, wherein the step (a) comprises determining the maximum buffer level close to a maximum buffer size when an

Inter frame to be encoded is near the previous Intra frame, and close to zero when the Inter frame is near the upcoming Intra frame.

5. The method as defined in any one of claims 1 to 4, wherein the step (c) comprises:

determining a difference between the maximum buffer level and the transmission buffer level; and

determining a margin of error between the target frame size and the maximum buffer size for encoding the Inter frames and the Intra frames based on the determined difference while keeping said margin of error.

6. The method as defined in any one of claims 1 to 5, wherein the step (a) comprises determining the maximum buffer level as follows:

$$\text{BufferLevel}_{MAX}(t) = \begin{cases} R_{MAX} & \text{if } t=0,1,\dots,L \\ \frac{R_{MAX}(FT-F-L)}{F(T-L-1)} + t \frac{R_{MAX}(1-F)}{F(T-F)} & \text{if } t = L+1,\dots,T-1 \end{cases}$$

wherein $\text{BufferLevel}_{MAX}(t)$ is the maximum buffer level, T is a period between two Intra frames, F is a frame rate, R_{MAX} is a maximum bitrate, L is a number of frames over which the transmission buffer level is optimized, and "t" is time.

7. The method as defined in any one of claims 1 to 6, wherein determining the target frame size comprises:

(a-1) determining a number of bits unused in encoding an Intra frame; and

(a-2) distributing the number of unused bits over the target frame size for N Inter frames following the Intra frame,

wherein N is a number of Inter frames over which a number of unused bits are distributed.

8. The method as defined in claim 7, wherein:

the step (a-1) comprises determining a difference between the encoded Intra frame size and a corresponding Intra frame target size; or

the step (a-2) comprises optimizing the transmission buffer level over a number L of Inter frames, where $0 \leq N \leq L$; or

the step (a-2) comprises distributing the number of unused bits substantially equally.

9. The method as defined in any one of claims 1 to 8, wherein the step (b) comprises determining the target frame size as follows:

$$B_p(t) = \begin{cases} B_p^I(t) & \text{if } t = 1, 2, 3, \dots, L \\ \frac{TR_{MAX} - FB_I - F \sum_{k=1}^L B_p^I(k)}{F(T - L - 1)} & \text{if } t = L, L+1, \dots, T-1 \end{cases}$$

where $B_p(t)$ is the target frame size of an Inter frame at time t , T is a period between two Intra frames, $B_p^I(t)$ is an Inter frame target size at time t , which is selected so that not to lead to overflowing the transmission buffer, B_I is an Intra frame size, F is a frame rate, k is an index in a summation function, L is a number of frames over which the transmission buffer level is optimized, and R_{MAX} is a maximum bitrate.

10. The method as defined in any one of claims 1 to 8, wherein the step (b) comprises determining the target frame size as follows:

$$B_p(t) = \begin{cases} \frac{R_{MAX}}{F} + \frac{(R_{MAX} - B_I)}{N} & \text{if } 0 < t \leq N \\ \frac{R_{MAX}}{F} & \text{if } N < t \leq L \end{cases}$$

$$B_p(t) = \frac{R_{MAX}(T-F-L)}{F(T-L-1)} \quad \text{if } t = L, L+1, \dots, T-1$$

where $B_p(t)$ is the target frame size of the Inter frame at time t , B_I is an Intra frame size, T is a period between two Intra frames, B_p^I is an Inter frame target size equal to R_{MAX}/F , F is a frame rate, L is a number of frames over which the transmission buffer level is optimized, and R_{MAX} is a maximum bitrate.

11. The method as defined in any one of claims 1 to 10, further comprising:

- (i) dropping the Inter frames when a size of the Inter frame after encoding exceeds the maximum buffer level; or
- (ii) storing respective encoder state variables before encoding the Intra frames or the Inter frames; or
- (iii) restoring the respective encoder state variables after an encoded Inter frame or an encoded Intra frame has been dropped prior to encoding a next Inter frame or a next Intra frame; or
- (iv) verifying that a size of the Inter frame or a size of the Intra frame after encoding is smaller than a given maximum limit, wherein the maximum limit is a maximum bitrate.

12. A device for rate controlling in video coding of sequences pictures, including a series of Inter frames separated by Intra frames, the device comprising:

- (a) means for determining, for an Inter frame of the series, a maximum buffer level of a transmission buffer as a function of a position of the Inter frame relative to a previous Intra frame and an upcoming Intra frame in the sequence of pictures;

- (b) means for determining a target frame size measured in bits for the Inter frame leading to a transmission buffer level not exceeding the maximum buffer level;
- (c) means for encoding the Inter frame into an encoded frame according to the target frame size, the encoded frame having an encoded frame size;
- (d) means for determining the transmission buffer level in response to the encoded frame; and
- (e) means for dropping the encoded frame if it would cause the maximum buffer level to be exceed.

13. The device as defined in claim 12, wherein the means (e) comprises means for restoring the transmission buffer level to a previous value, provided the encoded frame was dropped.

14. The device as defined in any one of claims 12 or 13, wherein the means (a) is configured:

to determine the maximum buffer level, which ensures that the transmission buffer can accommodate the upcoming Intra frame; or

to determine the maximum buffer level close to a maximum buffer size when an Inter frame to be encoded is near the previous Intra frame, and close to zero when the Inter frame is near the upcoming Intra frame; or

to determine the maximum buffer level as follows:

$$\begin{aligned}
 BufferLevel_{MAX}(t) &= R_{MAX} && \text{if } t=0,1,\dots,L \\
 BufferLevel_{MAX}(t) &= \frac{R_{MAX}(FT-F-L)}{F(T-L-1)} + t \frac{R_{MAX}(1-F)}{F(T-F)} && \text{if } t = L+1,\dots,T-1
 \end{aligned}$$

wherein $\text{BufferLevel}_{\text{MAX}}(t)$ is the maximum buffer level, T is a period between two Intra frames, F is a frame rate, R_{MAX} is a maximum bitrate, L is a number of frames over which the transmission buffer level is optimized, and "t" is time.

15. The device as defined in any one of claims 12 to 14, wherein the means (c) is configured:

to determine a difference between the maximum buffer level and the transmission buffer level; and

to determine a margin of error between the target frame size and the maximum buffer size for encoding the Inter frames and the Intra frames based on the determined difference while keeping said margin of error.

16. The device as defined in any one of claims 12 to 15, wherein the means (a) comprises:

(a-1) means for determining a number of bits unused in encoding an Intra frame; and

(a-2) means for distributing the number of unused bits over the target frame size for N Inter frames following the Intra frame,

wherein N is a number of Inter frames over which a number of unused bits are distributed.

17. The device as defined in claim 16, wherein:

the means (a-1) is configured to determine a difference between the encoded Intra frame size and a corresponding Intra frame target size; or

the means (a-2) is configured to optimize the transmission buffer level over a number L of Inter frames, where $0 \leq N \leq L$; or

the means (a-2) is configured to distribute the number of unused bits substantially equally.

18. The device as defined in any one of claims 12 to 15, wherein:

the means (b) is configured to determine the target frame size as follows:

$$B_p(t) = B_p^I(t) \quad \text{if } t = 1, 2, 3, \dots, L$$

$$B_p(t) = \frac{TR_{MAX} - FB_I - F \sum_{k=1}^L B_p^I(k)}{F(T - L - 1)} \quad \text{if } t = L, L+1, \dots, T-1$$

where $B_p(t)$ is the target frame size of an Inter frame at time t, T is a period

between two Intra frames, $B_p^I(t)$ is an Inter frame target size at time t, which is selected so that not to lead to overflowing the transmission buffer, B_I is an Intra frame size, F is a frame rate, k is an index in a summation function, L is a number of frames over which the transmission buffer level is optimized, and R_{MAX} is a maximum bitrate; or

the means (a) is configured to determine the target frame size as follows:

$$B_p(t) = \frac{R_{MAX}}{F} + \frac{(R_{MAX} - B_I)}{N} \quad \text{if } 0 < t \leq N$$

$$B_p(t) = B_p^I = \frac{R_{MAX}}{F} \quad \text{if } N < t \leq L$$

$$B_p(t) = \frac{R_{MAX}(T - F - L)}{F(T - L - 1)} \quad \text{if } t = L, L+1, \dots, T-1$$

where $B_p(t)$ is the target frame size of the Inter frame at time t, B_I is an Intra frame size, T is a period between two Intra frames, B_p^I is an Inter frame target size

equal to R_{max}/F , F is a frame rate, L is a number of frames over which the transmission buffer level is optimized, and R_{MAX} is a maximum bitrate.

19. The device as defined in any one of claims 12 to 18, the device being further configured:

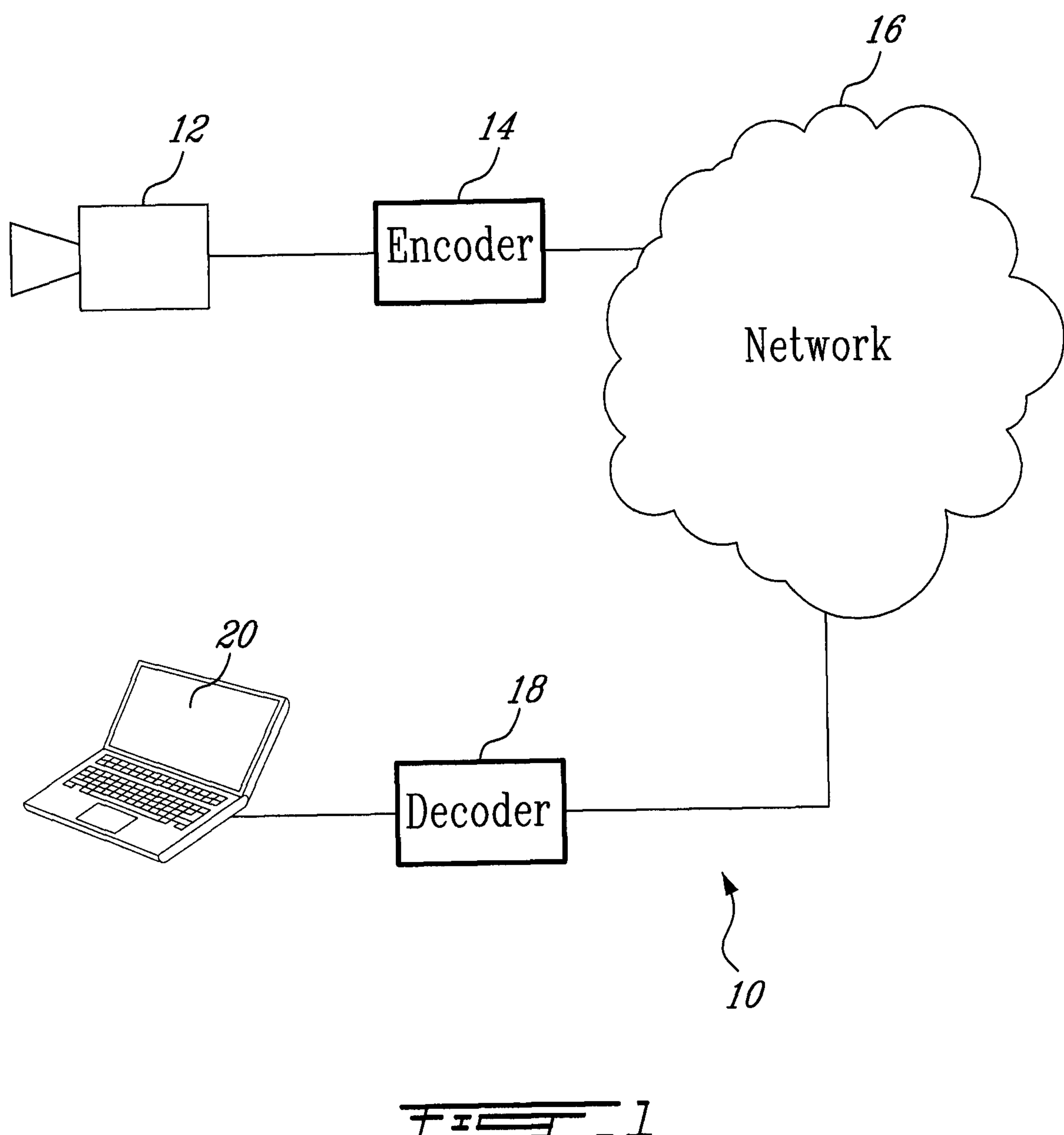
- (i) to drop the Inter frames when a size of the Inter frame after encoding exceeds the maximum buffer level; or
- (ii) to store respective encoder state variables before encoding the Intra frames or the Inter frames; or
- (iii) to restore the respective encoder state variables after an encoded Inter frame or an encoded Intra frame has been dropped prior to encoding a next Inter frame or a next Intra frame; or
- (iv) to verify that a size of the Inter frame or a size of the Intra frame after encoding is smaller than a given maximum limit, wherein the maximum limit is a maximum bitrate.

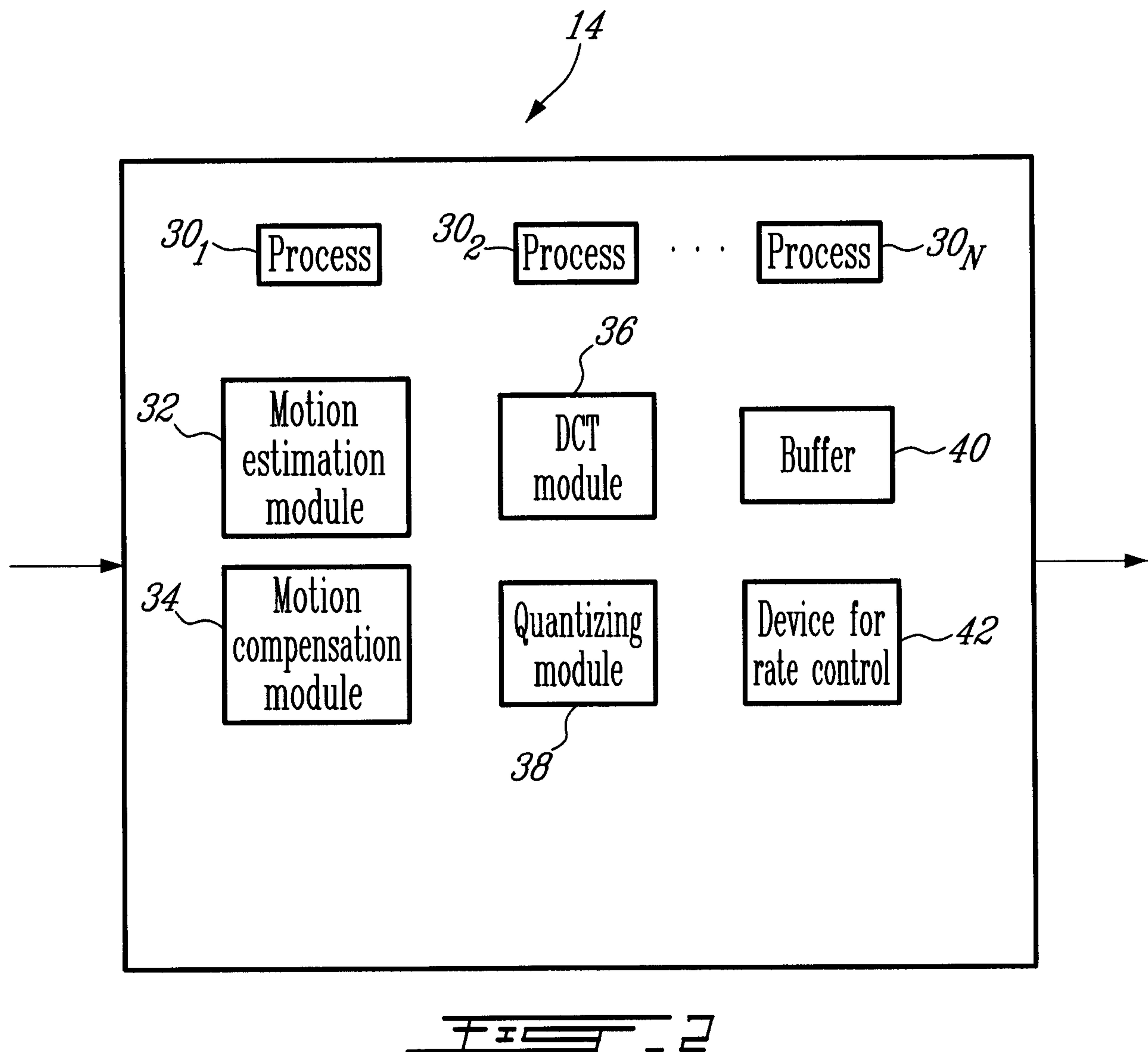
20. An encoder for video coding of sequences of pictures, including a series of Inter frames separated by Intra frames, the encoder comprising:

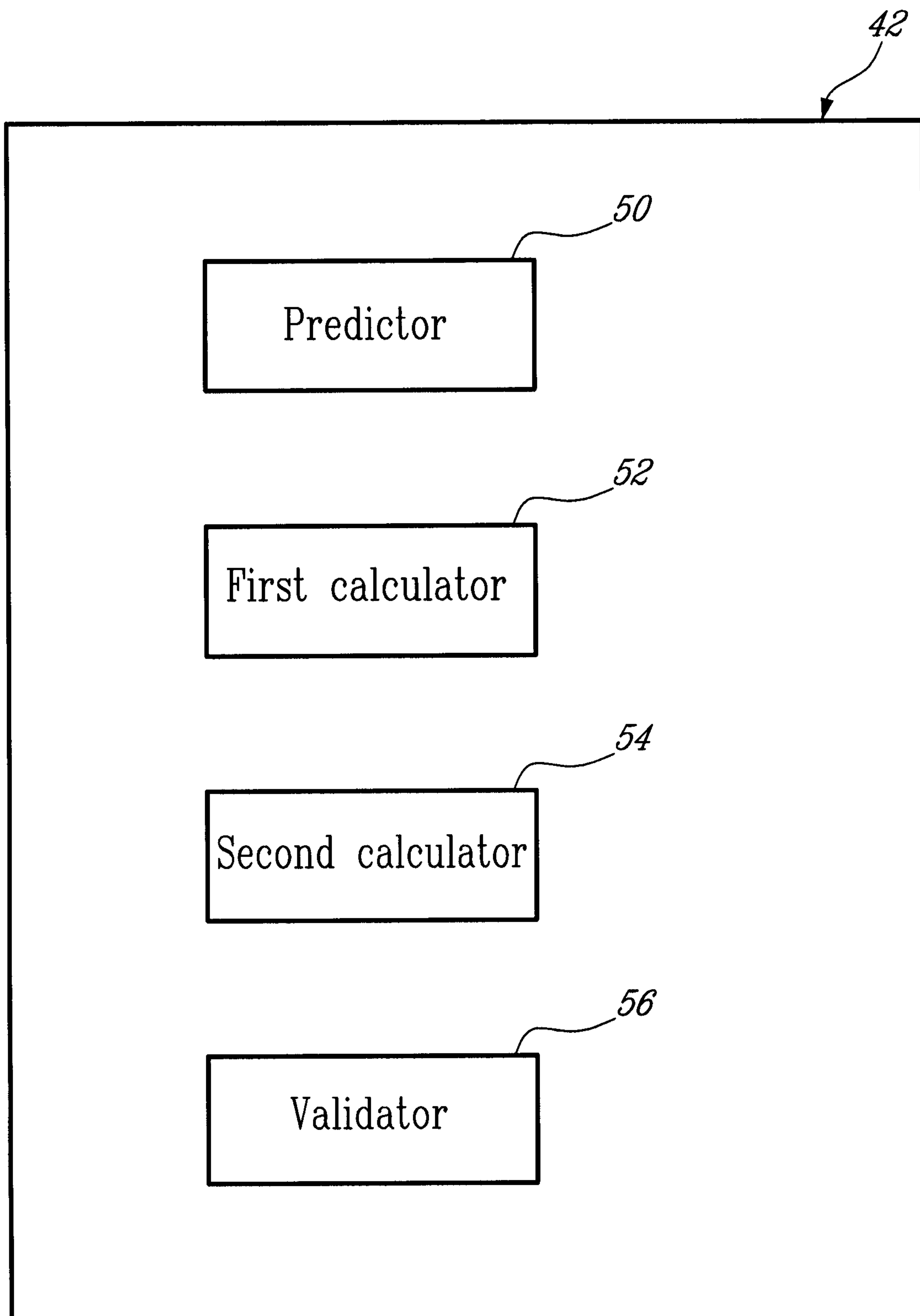
a processor; and

the device for rate controlling in the video coding as described in any one of claims 12 to 19.

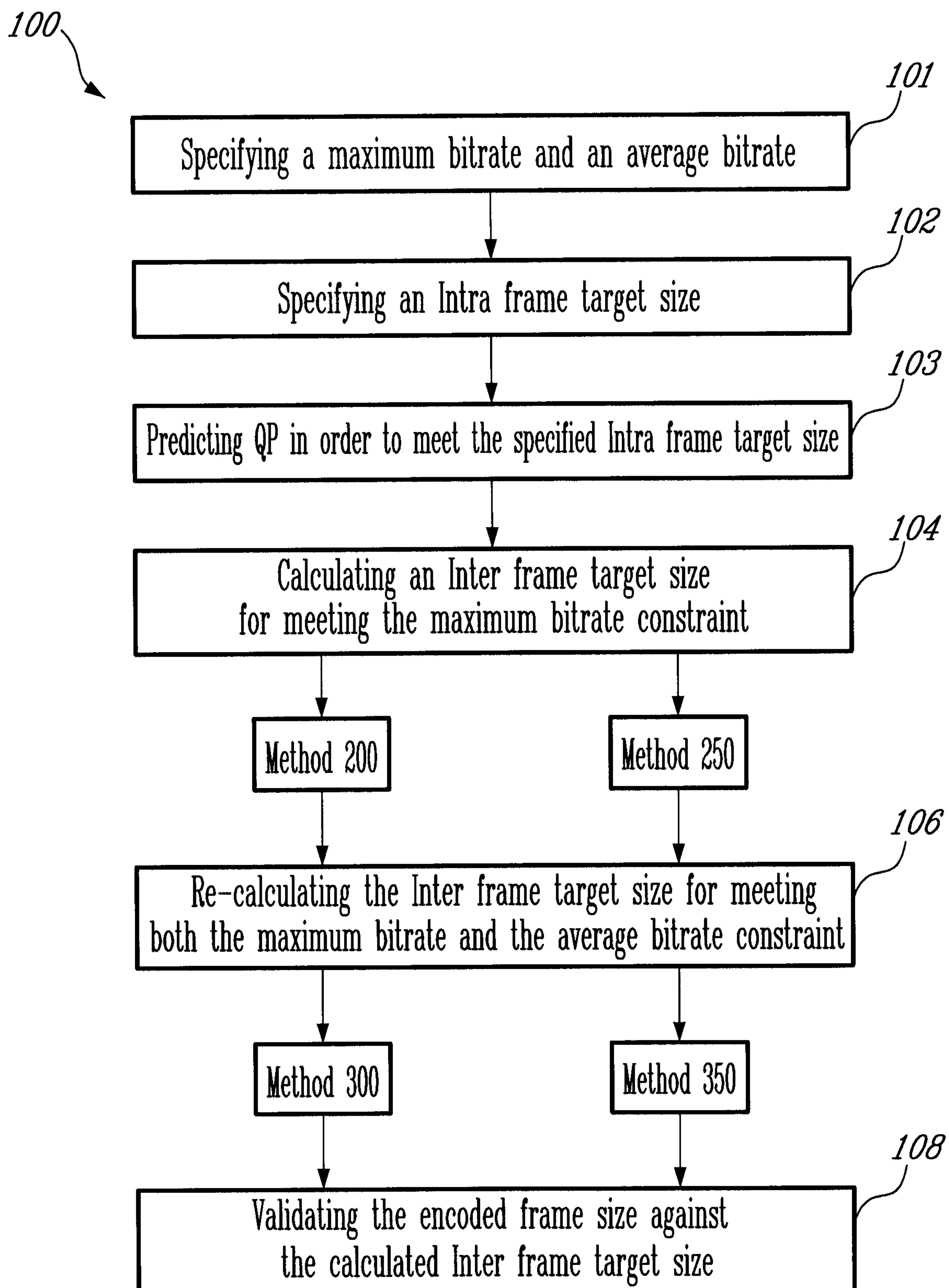
1 / 8







三 - 3



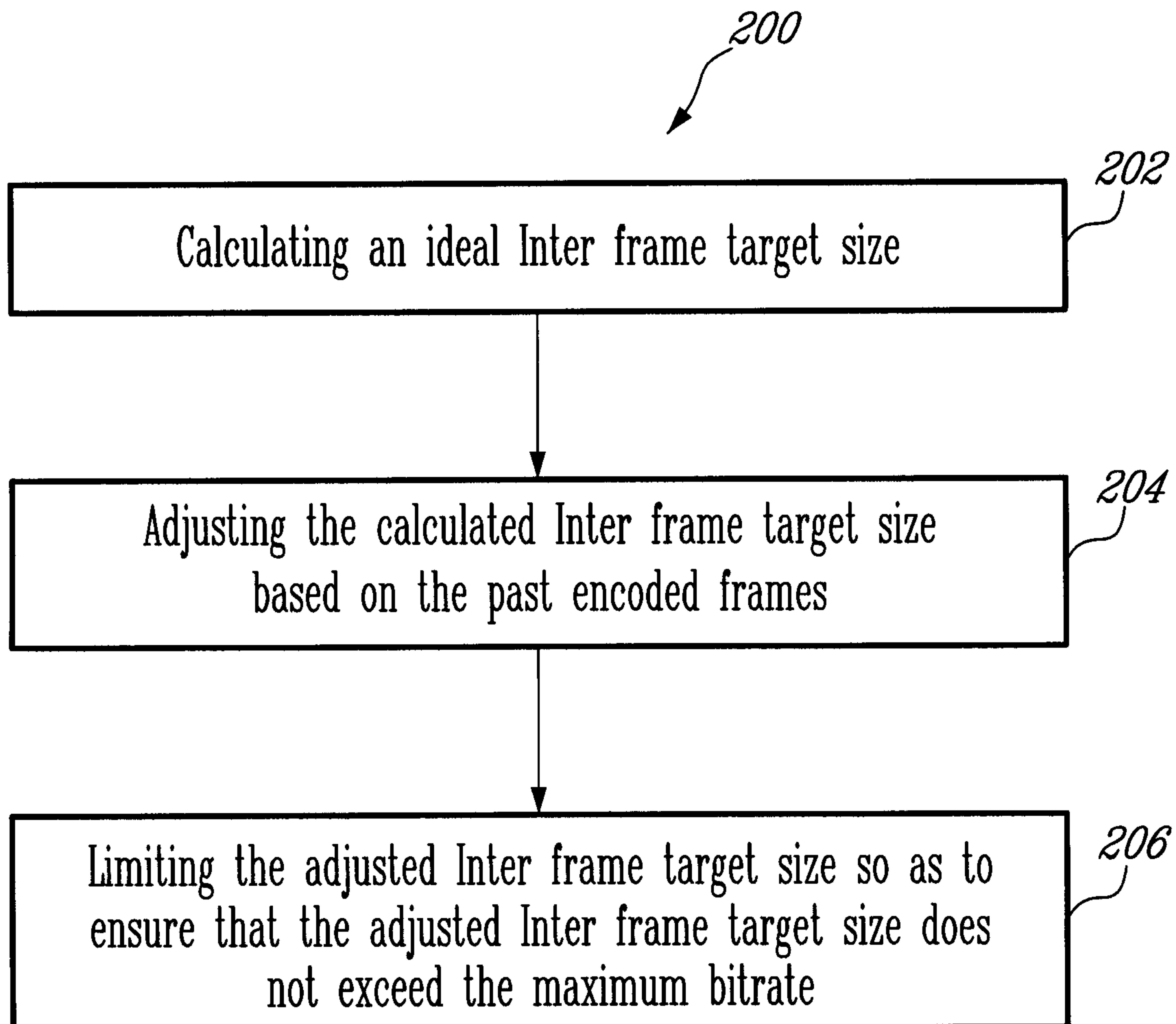


FIG - 5

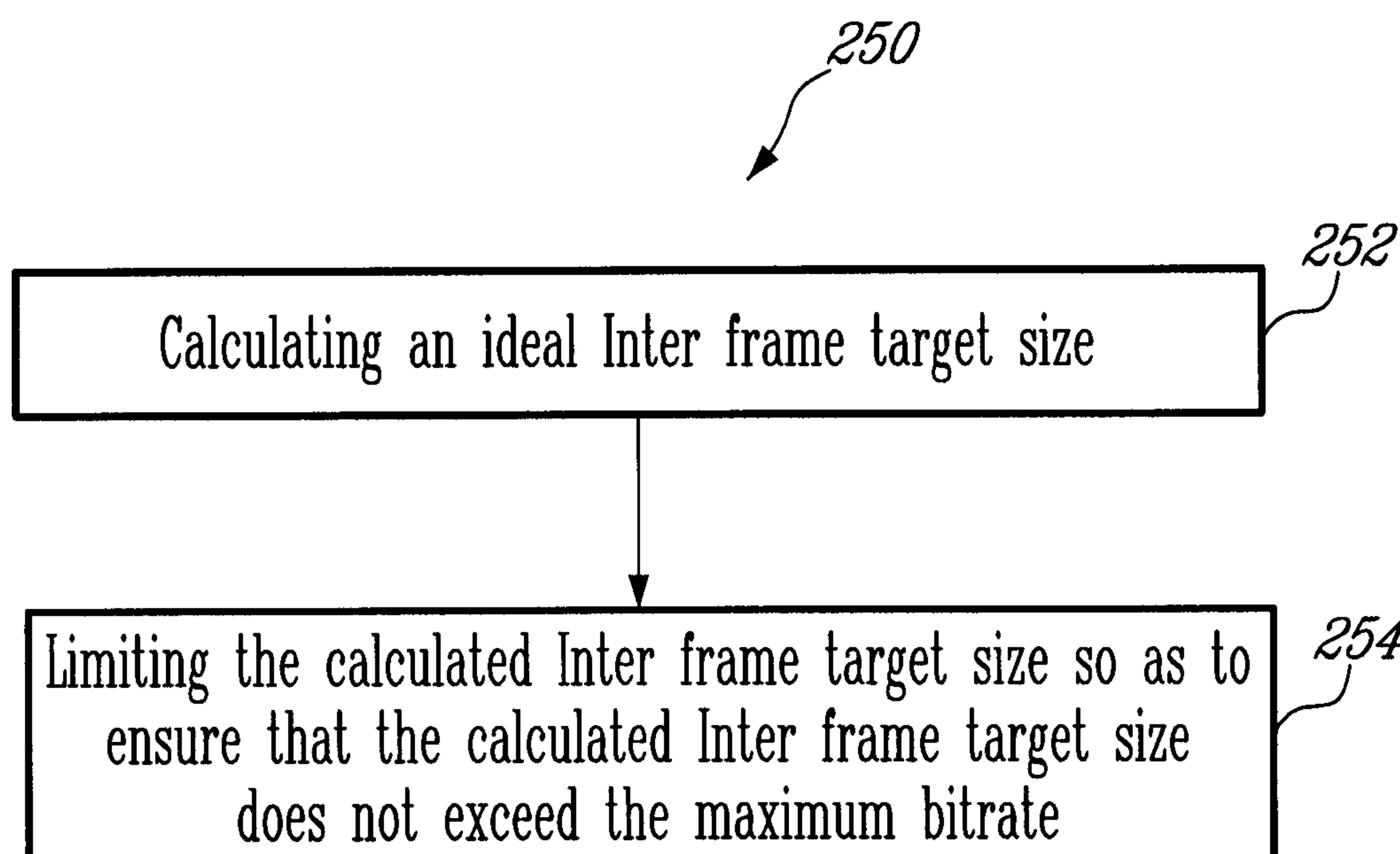
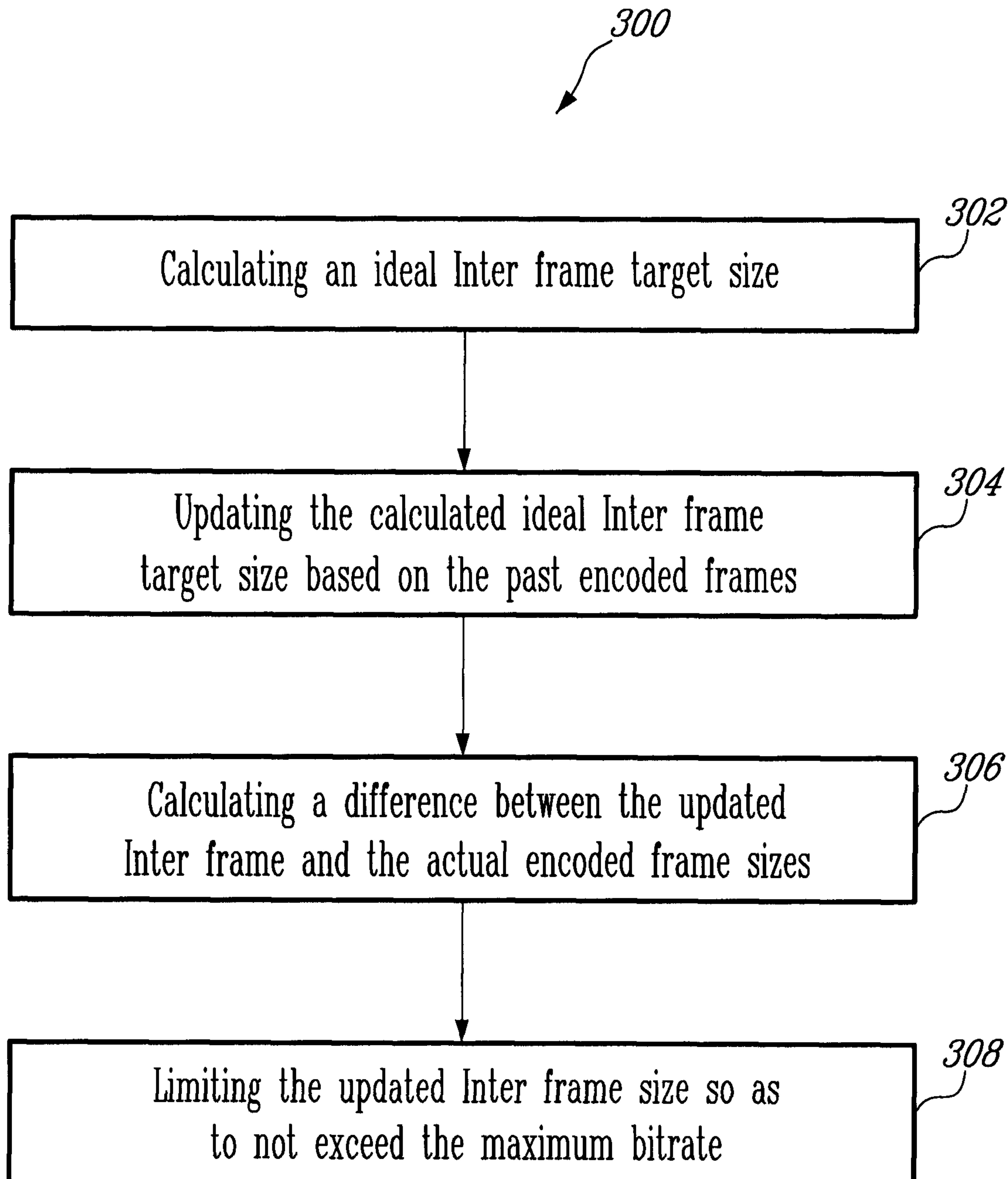


FIG - 6



— 7 —

