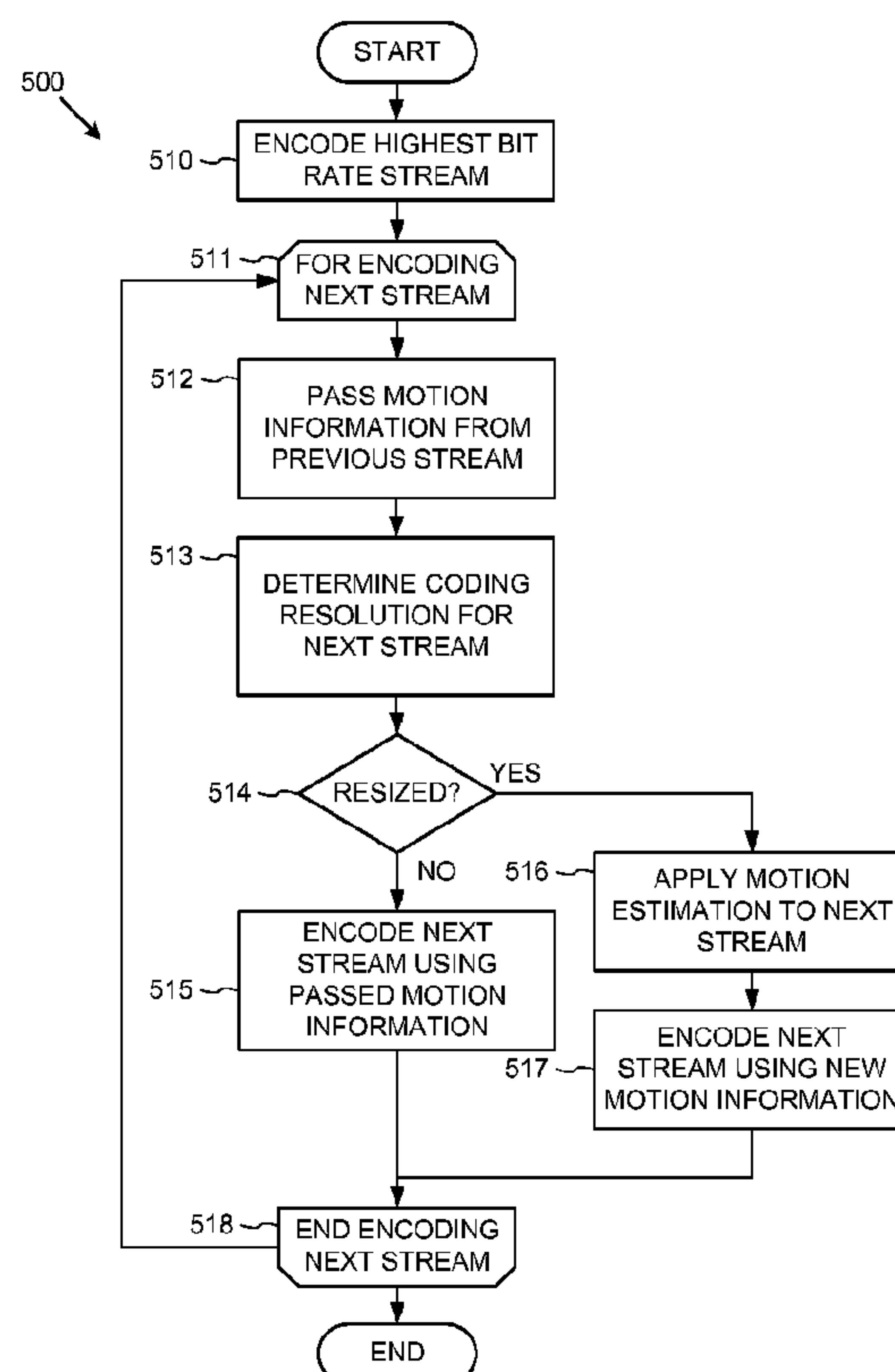




(86) Date de dépôt PCT/PCT Filing Date: 2010/01/12
(87) Date publication PCT/PCT Publication Date: 2010/08/05
(45) Date de délivrance/Issue Date: 2016/07/05
(85) Entrée phase nationale/National Entry: 2011/06/27
(86) N° demande PCT/PCT Application No.: US 2010/020789
(87) N° publication PCT/PCT Publication No.: 2010/088031
(30) Priorité/Priority: 2009/01/29 (US12/362,427)

(51) Cl.Int./Int.Cl. *H04N 7/24* (2011.01),
H04N 7/12 (2006.01)
(72) Inventeurs/Inventors:
GU, CHUANG, US;
CHAN, CHUN-WEI, US;
CHEN, WILLIAM, US;
SPEARS, STACEY, US;
HOLCOMB, THOMAS W., US
(73) Propriétaire/Owner:
MICROSOFT TECHNOLOGY LICENSING, LLC, US
(74) Agent: SMART & BIGGAR

(54) Titre : CODAGE VIDEO UTILISANT DES INFORMATIONS DE MOUVEMENT PRECEDEMMENT CALCULEES
(54) Title: VIDEO ENCODING USING PREVIOUSLY CALCULATED MOTION INFORMATION



(57) Abrégé/Abstract:

A video encoder uses previously calculated motion information for inter frame coding to achieve faster computation speed for video compression. In a multi bit rate application, motion information produced by motion estimation for inter frame coding of a compressed video bit stream at one bit rate is passed on to a subsequent encoding of the video at a lower bit rate. The video encoder chooses to use the previously calculated motion information for inter frame coding at the lower bit rate if the video resolution is unchanged. A multi core motion information precalculation produces motion information prior to encoding by dividing motion estimation of each inter frame to separate CPU cores.

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

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
5 August 2010 (05.08.2010)(10) International Publication Number
WO 2010/088031 A3(51) International Patent Classification:
H04N 7/24 (2006.01) *H04N 7/12* (2006.01)(21) International Application Number:
PCT/US2010/020789(22) International Filing Date:
12 January 2010 (12.01.2010)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
12/362,427 29 January 2009 (29.01.2009) US(71) Applicant (for all designated States except US): **MICROSOFT CORPORATION** [US/US]; One Microsoft Way, Redmond, Washington 98052-6399 (US).(72) Inventors: **GU, Chuang**; One Microsoft Way, Redmond, Washington 98052-6399 (US). **CHAN, Chun-Wei**; One Microsoft Way, Redmond, Washington 98052-6399 (US). **CHEN, William**; One Microsoft Way, Redmond, Washington 98052-6399 (US). **SPEARS, Stacey**; One Microsoft Way, Redmond, Washington 98052-6399 (US).**HOLCOMB, Thomas, W.**; One Microsoft Way, Redmond, Washington 98052-6399 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: VIDEO ENCODING USING PREVIOUSLY CALCULATED MOTION INFORMATION

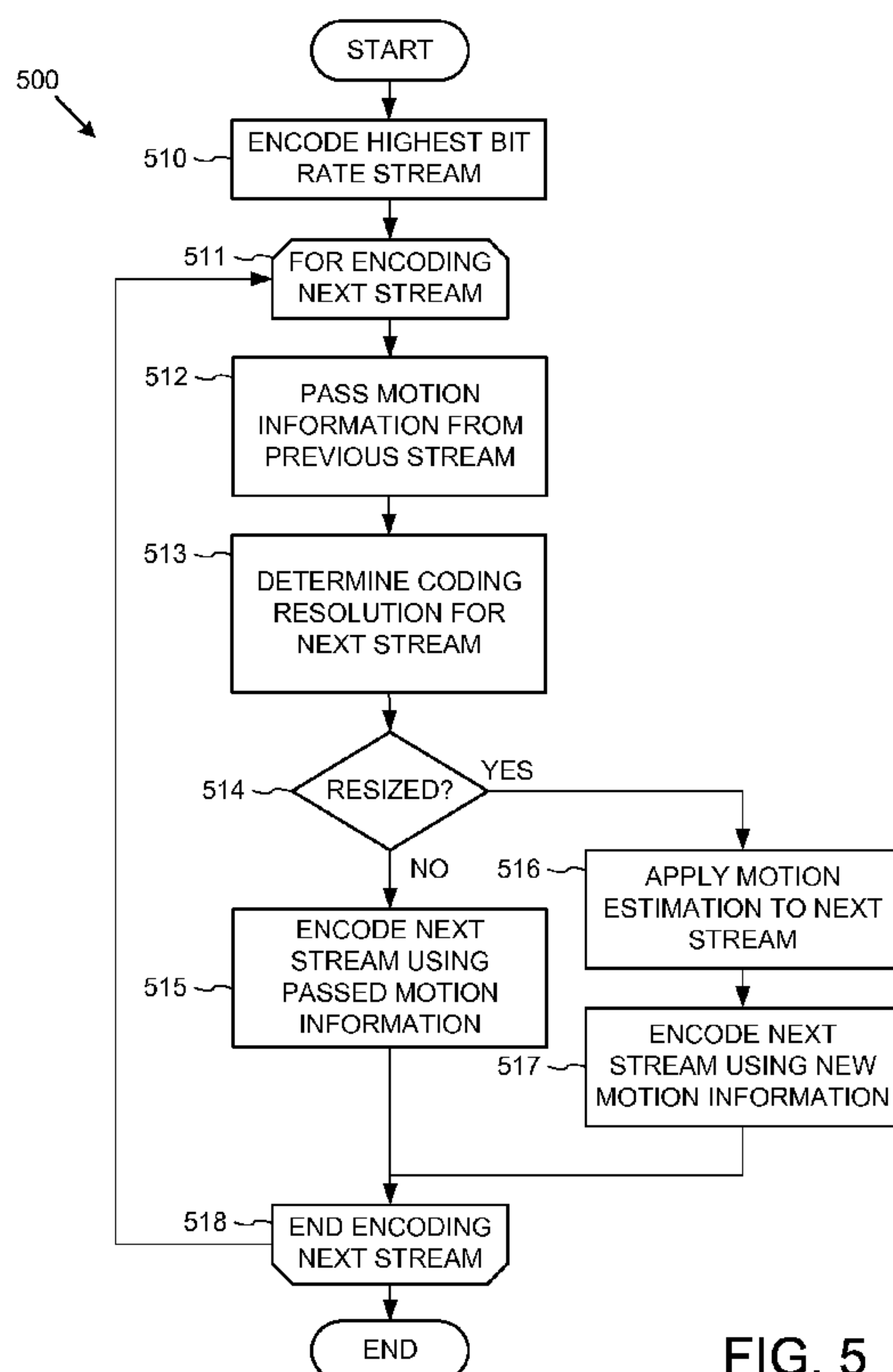


FIG. 5

(57) Abstract: A video encoder uses previously calculated motion information for inter frame coding to achieve faster computation speed for video compression. In a multi bit rate application, motion information produced by motion estimation for inter frame coding of a compressed video bit stream at one bit rate is passed on to a subsequent encoding of the video at a lower bit rate. The video encoder chooses to use the previously calculated motion information for inter frame coding at the lower bit rate if the video resolution is unchanged. A multi core motion information pre-calculation produces motion information prior to encoding by dividing motion estimation of each inter frame to separate CPU cores.

WO 2010/088031 A3



Declarations under Rule 4.17:

- | | | | |
|---|--|------|---|
| — | <i>as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))</i> | — | <i>before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))</i> |
| — | <i>as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))</i> | (88) | Date of publication of the international search report:
21 October 2010 |

Published:

- *with international search report (Art. 21(3))*

VIDEO ENCODING USING PREVIOUSLY CALCULATED MOTION INFORMATION

Background

[001] With the increasing popularity of playing streaming audio and video over networks
5 such as the internet, there is a need for optimizing the data transferred from a server to a
client such that the client's experience is maximized even if network conditions during
playback are inconsistent. Optimizing the client's experience involves choosing a quality
level for encoding the audio and video portions of the video playback such that the video
can be transferred and reconstructed uninterrupted while preserving the quality of the
10 video content.

[002] The quality level is generally dictated by the bit rate specified for the encoded
audio or video portions of the input stream. A higher bit rate generally indicates that a
larger amount of information about the original audio or video is encoded and retained,
and therefore a more accurate reproduction of the original input audio or video will be
15 presented during video playback. Conversely, a lower bit rate indicates that less
information about the original input audio or video is encoded and retained, and thus a less
accurate reproduction of the original audio or video will be presented during video
playback.

[003] Generally, the bit rate is specified for encoding each of the audio and video based
20 on several factors. The first factor is the network condition between the server and the
client. A network connection that can transfer a high amount of data indicates that a
higher bit rate can be specified for the input video that is subsequently transferred over the
network connection. The second factor is the desired start-up latency. Start-up latency is
the delay that a video playback tool experiences when first starting up due to the large
25 amount of data that has to be received, processed, and buffered. The third factor is the
tolerance to glitching. Glitching is when video playback has to stop because data is
missing. In most cases any amount of start-up latency or glitching is intolerable, and it is
therefore desirable to optimize the bit rate specified such that the start-up latency and the
glitching are minimized or eliminated.

30 [004] Currently available commercial streaming media systems rely on multi bit rate
(MBR) coding to perform coding rate control. In MBR coding, source video content is
encoded into alternative bit streams at different coding rates and typically stored in the
same media file at the server. This then allows the content to be streamed in segments or

chunks at varying levels of quality corresponding to different coding rates according to the changing network conditions, typically using bit stream switching between segments.

Motion Estimation and Compensation in Video Encoding

[005] Numerous techniques have been developed for encoding video to be compressed to a desired bit rate. Such techniques include intraframe compression techniques (in which a frame is compressed as a still image) and interframe compression techniques (in which a frame is predicted or estimated from one or more other frames). Intraframe compression often involves frequency transformations on data followed by lossy and lossless compression. Interframe compression can include motion estimation.

[006] Motion estimation is a process for estimating motion between frames. Motion estimation for video is said to be an ill-posed problem, because the motion within the viewed scene occurs in 3 dimensions, but the successive frames of video are a projection of the 3D scene onto a 2 dimensional image plane. In one common technique, an encoder using motion estimation attempts to match a block of pixels in a current frame with a similar block of pixels (picture elements or samples) in a search area in another frame (called the reference frame). When the encoder finds an exact or "close enough" match in the search area in the reference frame, the encoder parameterizes the change in position of the blocks as motion data (such as a motion vector).

[007] Conversely, motion compensation is a process of reconstructing frames from reference frames using motion data. In one common technique, an encoder or decoder reconstructs a current frame by applying motion data for the current frame to a reference frame, creating a predicted frame. The encoder can compress the difference (sometimes called the residual) between the predicted frame and the original version of the current frame using the same techniques as used for intraframe compression (e.g., lossy and lossless compression). The overall bit rate of the camera video depends very much on the bit rate of the residuals, which can predominate in the overall bit rate compared to the bit rate for motion data. The bit rate of residuals is low if the residuals are simple (i.e., due to motion estimation that leads to exact or good matches according to some criteria), or if lossy compression drastically reduces the complexity of the residuals. On the other hand, the bit rate of complex residuals (i.e., those for which motion estimation fails to find good matches) can be higher, depending on the degree of lossy compression applied to reduce the complexity of the residuals.

Summary

[008] The following Detailed Description concerns techniques (implemented via methods, devices and systems) for video encoding using previously calculated motion information, such as in multi bit rate and multi-core applications. In these various applications, a video encoder takes both raw video content and previously calculated motion information as input. The video encoder can then apply motion compensation based on the previously calculated motion information to encode predicted frames of the raw video content. For various applications, this allows the video encoder to more efficiently apply motion estimation and compensation.

[009] In an example multi bit rate application, the video encoder with previously calculated motion information input is used to encode source video into a plurality of bit streams at successively lower bit rates. For each bit stream encoded at the successive lower bit rate, the video encoder receives the motion information calculated for the previous bit stream at the higher bit rate as its previously calculated motion information input. The video encoder can choose to use this input motion information calculated for the previous bit stream at the higher bit rate to perform motion compensation encoding of the bit stream at a next lower bit rate. However, the video encoder will need to make various adjustments of the encoding parameters to achieve the lower bit rate, such as degree of quantization and video resolution, among other encoding parameters. Because motion estimation for video is an “ill-posed problem” as discussed in the Background above (among other reasons), the motion information previously calculated for the video at the higher encoding bit rate may prove inaccurate for encoding the same source video with the encoding parameters adjusted to achieve the lower bit rate. For example, the motion vector that yields a closer match (with smaller residual) for a given block of a predicted frame can vary between when the reference frame is encoded with the higher bit rate encoding parameters, as compared to when the reference frame is encoded at the lower bit rate. In the example multi bit rate application, the video encoder chooses whether to use the input motion information calculated for the higher bit rate depending on whether the encoding for the next successive bit rate includes a change of video resolution. In the case that the video resolution has not changed for the next successive bit rate, the video encoder uses the input motion information previously calculated for the higher bit rate. But if the video resolution is changed, the video encoder instead calculates the motion information anew for encoding the bit stream at the current bit rate. This newly

51331-1055

calculated motion information is passed on as the previously calculated motion information input to the video encoder for the bit stream at the next lower bit rate.

[010] In an example multi core processor application, the motion estimation is performed as a separate pre-calculation phase before input to the video encoder.

- 5 Intraframe compression is first performed on reference frames of the source video. The motion estimation pre-calculation can then be performed as a largely parallel operation on a multi core computer. Each processor core of the computer is assigned to calculate the motion estimation of a separate frame that is to be encoded using inter-frame (predicted) encoding relative to its respective reference frame. With
10 the computation of motion estimation split in this manner between the processor cores, the video encoder is able to more fully utilize the processing capacity of the multi core computer.

- [010a]** According to one aspect of the present invitation, there is provided a method of processing raw video content input for encoding into multi bit rate video,
15 the method comprising: receiving raw video content for encoding as compressed video bit streams at a succession of progressively lower bit rates; encoding an initial compressed video bit stream at a highest bit rate, said encoding the initial compressed video bit stream producing motion information for encoding of inter frame pictures of the video; for encoding subsequent compressed video bit streams
20 at the progressively lower bit rates, passing the motion information produced from encoding a preceding compressed video bit stream to the encoding of the subsequent compressed video bit stream; determining whether a condition based on the video encoding parameters for the subsequent compressed video bit stream is met; if said condition is met, encoding the subsequent compressed video bit stream
25 wherein inter frames of the subsequent compressed video bit stream are encoded using the motion information passed from encoding the preceding compressed video bit stream; and if said condition is not met, producing new motion information by performing motion estimation based on reconstructed reference frames of the subsequent compressed video bit stream, and encoding the subsequent compressed

51331-1055

video bit stream wherein inter frames of the subsequent compressed video bit stream are encoded using the new motion information; and producing an output multi bit rate video comprising the initial and subsequent compressed video bit streams.

[010b] According to another aspect of the present invention, there is provided

5 a method of encoding video on a computer having a multi core CPU, the method comprising: producing motion information for groups of inter frame pictures of a video as a motion estimation pre-calculation prior to encoding of the video, wherein said producing motion information comprises: for each inter frame picture in a group, performing motion estimation on a separate one of plural CPU cores of the computer,
10 said motion estimation producing motion information for plural sets of elements of the respective inter frame picture; encoding the video, wherein said encoding comprises encoding the inter frame pictures based on the motion information produced by the motion estimation pre-calculation; and producing an output compressed video bit stream containing the encoded video.

15 **[010c]** According to still another aspect of the present invention, there is

provided a video processing system for encoding compressed video streams for multiple bit rate video streaming, the system comprising: a memory for storing raw video content to be encoded; a computer processing means programmed for encoding the raw video content as a plurality of compressed video streams, wherein
20 said encoding the compressed video streams comprises: producing original motion information for encoding inter-frame pictures of the raw video content via motion estimation; encoding a top quality compressed video bit stream from the raw video content using the original motion information for encoding inter-frame pictures of the raw video content; deciding a video resolution for encoding a subsequent
25 compressed video bit stream at a lower bit rate; if the video resolution for encoding the subsequent compressed video bit stream is unchanged, encoding the subsequent compressed video bit stream from the raw video content again using the original motion information for encoding inter-frame pictures of the raw video content; and
30 otherwise if the video resolution for encoding the subsequent compressed video bit stream is resized from that of the top quality compressed video bit stream, then

51017-28

producing new motion information for encoding inter-frame pictures of the raw video content via motion estimation relative to intra-frame reference pictures of the subsequent compressed video bit stream, and encoding the subsequent compressed video bit streams from the raw video content using the new motion information for encoding inter-frame pictures of the raw video content.

[010d] According to yet another aspect of the present invention, there is provided a computer readable medium having computer executable instructions stored thereon for execution by one or more computers, that when executed implement a method as described above, or detailed below.

[011] This Summary is provided to introduce a selection of concepts in a simplified form that is further described below in the Detailed Description. This summary is not intended to identify key features or essential features of the claimed subject matter, nor is it intended to be used as an aid in determining the scope of the claimed subject matter. Additional features and advantages of the invention will be made apparent from the following detailed description of embodiments that proceeds with reference to the accompanying drawings.

Brief Description Of The Drawings

[012] Figure 1 is a block diagram of a video encoder with precalculated motion information input.

[013] Figure 2 is a block diagram of a generalized implementation of the video encoder of Figure 1.

[014] Figure 3 is a block diagram of a video streaming system that provides segmented streaming of video at variable bit rates.

[015] Figure 4 is a block diagram of a video encoding system for encoding a segmented video streaming file using the video encoder of Figure 1 for use in segmented streaming by the video streaming system of Figure 3.

51017-28

[016] Figure 5 is a flow chart of the process performed by the video encoding system of Figure 4.

[017] Figure 6 is a data flow diagram of a video encoding system that uses multi core processors for pre-calculation of motion information for the video encoder of Figure 1.

5 [018] Figure 7 is a block diagram of a generalized operating environment in conjunction with which various described embodiments may be implemented.

Detailed Description

[019] The following detailed description concerns various techniques and systems for video encoding using previously calculated motion information. The techniques described in the context of a multi bit rate streaming application, and a multi core video encoding application. However, the techniques can be applied more broadly to other video encoding applications.

[020] The various techniques and tools described herein may be used independently. Some of the techniques and tools may be used in combination. Various techniques are described below with reference to flowcharts of processing acts. The various processing acts shown in the flowcharts may be consolidated into fewer acts or separated into more acts. For the sake of simplicity, the relation of acts shown in a particular flowchart to acts described elsewhere is often not shown. In many cases, the acts in a flowchart can be reordered.

I. Video Encoder Using Precalculated Motion Information

[021] Figure 1 depicts one example of a video encoder 100 that can be used in the multi bit rate and multi core applications described below. The video encoder 100 has inputs 110, 120 for receiving “raw” (uncompressed) frames of video content and also previously calculated motion information for the video content. The video encoder then performs intra-frame coding of reference frames of the video content, and utilizes the motion information to perform inter-frame coding of the predicted frames of the video content. The encoding can be performed according to a known video encoding standard, such as Windows Media Video format, SMPTE 421-M format, MPEG-x format (e.g., MPEG-1, MPEG-2, or MPEG-4), H.26x format (e.g., H.261, H.262, H.263, or H.264), or other format. However, in the case of inter-frame coding, the video encoder can choose to use the pre-calculated motion information for the inter-frame coding of a predicted frame, rather than performing its own motion estimation for the frame. The video encoder encodes the video content into a compressed bitstream provided as output 130. The video encoder may also output the motion information that it used for inter-frame compression of the input video content as motion information output 140.

[022] Figure 2 is a generalized block diagram showing one example suitable implementation of the video encoder 100, in conjunction with which the multi bit rate and multi core techniques described more fully below may be implemented. The video encoder 100 receives a sequence of video pictures (frames) as its raw video content input

110 along with pre-calculated motion information input 120 and produces a compressed bit stream 130 as output.

[023] The video encoder 100 processes video pictures. The term “picture” generally refers to source, coded, or reconstructed image data. For progressive video, a picture is a progressive video frame. For interlaced video, a picture may refer to an interlaced video frame, the top field of the frame, or the bottom field of the frame, depending on context.

[024] The video encoder 100 compresses inter-coded, predicted pictures of the input video and intra-coded pictures of the input video. For the sake of presentation, Figure 2 shows a path for intra-coded frames through the encoder 100 and a path for inter-coded predicted frames. Many of the components of the video encoder 100 are used for compressing both intra-coded content and inter-coded, predicted content. The exact operations performed by those components can vary depending on the type of information being compressed.

[025] In general, within the video encoder 100, an inter-coded, predicted frame (as a picture) is represented in terms of prediction from previously reconstructed content (as one or more other pictures, which are typically referred to as reference pictures or anchors). For example, content at a given time is encoded as a progressive P-frame or B-frame, interlaced P-field or B-field, or interlaced P-frame or B-frame. Within the video encoder 100, a prediction residual is the difference between predicted information and corresponding intra-coded frames.

[026] The input video 110 content on the inter-path is encoded as a predicted picture based on motion information. If certain conditions are met, the video encoder 100 uses the pre-calculated motion information from input 120 (as illustrated by selection switch 256), which can be in the form of a set or sequence of motion vector for macroblocks or other sets of samples of the inter-path video picture with respect to one or more reference pictures. In general, the choice to use the pre-calculated motion information can be based on: first, the availability of pre-calculated motion information; and second, which and whether encoding parameters were changed from the previous calculation of the motion information and the parameters used for the current encoding of the video content. In one example, the video encoder will choose not to use the previously calculated motion information from input 130 if the motion information was calculated for encoding the video content with a different video resolution than that which the video encoder is current encoding.

[027] However, the video encoder 100 can instead choose (again illustrated by selection switch 256) to perform new motion estimation for the inter-path video content 110 with motion estimator 258. The motion estimator 258 estimates motion of macroblocks or other sets of samples of the video picture with respect to one or more reference pictures, which represent reconstructions of previously encoded video content frames. The picture store 264 buffers this reconstructed video content 266 as a reference picture or pictures. When multiple reference pictures are used, the multiple reference pictures can be from different temporal directions or the same temporal direction. The motion estimator 258 outputs motion information 260 such as motion vector information.

[028] The motion compensator 262 applies motion vectors (whether from the pre-calculated motion information input or from new motion estimation) to certain reconstructed video content 266 (stored as reference picture(s)) when forming a motion-compensated current picture 268. The difference (if any) between a block of the motion-compensated picture 268 and corresponding block of the original inter-path video picture is the prediction residual 270 for the block. During later reconstruction of the inter-path video frame (e.g., at a video decoder), reconstructed prediction residuals are added to the motion compensated residual video 268 to obtain reconstructed content closer to the original inter-path video 256. In lossy compression, however, some information is still lost from the original inter-path video. Alternatively, a motion estimator and motion compensator apply another type of motion estimation/compensation.

[029] A frequency transformer 280 converts spatial domain video information into frequency domain (i.e., spectral, transform) data. For block-based video content, the frequency transformer 280 applies a DCT, variant of DCT, or other forward block transform to blocks of the samples or prediction residual data, producing blocks of frequency transform coefficients. The frequency transformer 280 may apply an 8x8, 8x4, 4x8, 4x4 or other size frequency transform.

[030] A quantizer 282 then quantizes the blocks of transform coefficients. The quantizer 282 applies non-uniform, scalar quantization to the spectral data with a step size that varies spatially on a picture-by-picture basis, macroblock-by-macroblock basis or other basis. Additionally, in some cases the quantizer varies quantization across color channels of the inter-layer residual video picture. The quantizer 282 can also apply another type of quantization, for example, a uniform or adaptive quantization for at least some spectral data coefficients, or directly quantizes spatial domain data in an encoder system that does not use frequency transformations.

[031] When reconstructed video content is needed for subsequent motion estimation/compensation of an inter-path video picture, an inverse quantizer 290 performs inverse quantization on the quantized spectral data coefficients. An inverse frequency transformer 292 performs an inverse frequency transform, producing blocks of reconstructed prediction residuals (for predicted inter-path residual video content) or samples (for intra-path residual video content). If the residual video content 256 was motion-compensation predicted, the reconstructed prediction residuals are added to the motion-compensated predictors 268 to form the reconstructed residual video. The picture store 264 buffers the reconstructed residual video for use in subsequent motion-compensated prediction.

[032] The entropy coder 284 compresses the output of the quantizer 282 as well as certain side information (e.g., quantization parameter values). Typical entropy coding techniques include arithmetic coding, differential coding, Huffman coding, run length coding, LZ coding, dictionary coding, and combinations of the above. The entropy coder 284 typically uses different coding techniques for different kinds of information, and can choose from among multiple code tables within a particular coding technique.

[033] When the video encoder 240 performs intra-compression of the intra-path video content, the encoder intra-compresses it as an intra-coded picture, without motion compensation. The video 256 is provided directly to the frequency transformer 280, quantizer 282, and entropy coder 284 and output as encoded video. A reconstructed version of the intra-coded video can be buffered for use in subsequent motion compensation of other inter-path video.

[034] A controller 294 receives inputs from various modules such as the motion estimator 258, frequency transformer 280, quantizer 282, inverse quantizer 290, and entropy coder 284. The controller 294 evaluates intermediate results during encoding, for example, setting quantization step sizes and performing rate-distortion analysis. The controller 294 works with other modules to set and change coding parameters during encoding. When the controller 294 evaluates different coding parameter choices, the controller 294 may iteratively perform certain stages to evaluate different parameter settings, or the controller 294 may jointly evaluate different coding parameters. The tree of coding parameter decisions to be evaluated, and the timing of corresponding encoding, depends on implementation. In some embodiments, the controller 294 also receives input from an encoding session wizard interface, other encoder application interface, or other source to designate video to be encoded using specific rules.

[035] In certain applications, the video encoder 100 also may output the motion information that was used for motion compensation of the inter-path video content (whether from pre-calculated motion information input 130 or motion estimator 258), such as for use as pre-calculated motion information for encoding the raw video content using
5 varying other encoding parameters to achieve a different bit rate in the multi bit rate application discussed below.

II. Multi Bit Rate Video Encoder

[036] The video encoder 100 illustrated in Figures 1 and 2 can be used to produce an indexed file for multi bit rate video streaming applications. The indexed file generally
10 divides video of a multimedia program into multiple streaming segments, and contains a number of compressed bit streams representing the video segments at various bit rates.

[037] In general, the indexed file for multi bit rate streaming can be used by standard HTTP servers to serve multimedia content at multiple bit rates with bit rate selection (rate control) being performed client-side (e.g., exclusively client-side). Clients can perform
15 rate control by first obtaining index information from the server describing the various bit rates available for streaming segments of a program. Based on the index information, and possibly other information (e.g., network bandwidth, buffer information, etc.), the client can decide which bit rate streaming segments to download from the server to provide a desired user experience (e.g., the best user experience possible based on the available bit
20 rates and current network conditions).

[038] Other types of computing devices (e.g., other than traditional HTTP servers) can provide files using the indexed file. For example, a computing device (e.g., a personal computer, server computer, or special-purpose streaming media server) can use the indexed file layout to serve multimedia content using various file serving protocols (e.g.,
25 File Transfer Protocol (FTP), Hypertext Transfer Protocol (HTTP), Real Time Streaming Protocol (RTSP), MMS (Microsoft Media Services), etc.).

[039] In order to support bit rate switching, programs are divided into temporal chunks called streaming segments (self-contained units). The server stores each streaming segment at one or more bit rates (e.g., each streaming segment – bit rate combination is a
30 separate streaming segment encoding). Each streaming segment includes one or more available bit rate encodings for a specific track (e.g., a specific audio track, such as an English audio track, or a specific video track) of a program. Clients then determine which bit rate, from the available bit rates (e.g., from the available streaming segment encodings), to download for each streaming segment. For example, a client may obtain a

first streaming segment, for a video track, encoded at 250Kb/sec (kilo-bits per second) (from one or more available streaming segment encodings for the first streaming segment), a second streaming segment, for the video track, encoded at 500Kb/sec (from one or more available streaming segment encodings for the second streaming segment), and a third
5 streaming segment, for the video track, encoded at 1Mb/sec (mega-bit per second) (from one or more available streaming segment encodings for the third streaming segment). Each streaming segment encoding can be at a constant bit rate (CBR) or at a variable bit rate (VBR) (and the use of CBR and VBR can be mixed on a per streaming segment encoding basis).

10 **[040]** Figure 3 depicts a generalized block diagram 300 of a system for segmented streaming of multimedia content using an indexed file layout. In the diagram 300, a server 310 (e.g., a server computer system such as a standard HTTP server) provides multimedia content to a client 320 (e.g., a client computer system, such as a laptop or desktop
15 computer, or another type of computing device, such as a PDA or mobile phone) via a network 330 (e.g., the Internet). In the diagram 300, the server 310 stores programs in an indexed file. The client 320 comprises client-side rate control software and/or hardware.

[041] In one specific example implementation, the server 310 is a standard HTTP server without any specialized streaming capability other than the ability to serve files. Because the server 310 does not support any specialized bit rate selection capability, the client 320
20 must perform all bit rate selection activities. In this implementation, the client 320 performs all bit rate selection activities. For example, the client 320 can perform rate control using the index information obtained from the server 310 (e.g., alone or in combination with other information, such as client buffer information, network bandwidth, etc.). However, in other implementations, some or all of the rate-control functions can
25 occur at the server.

[042] Figure 4 illustrates a multi bit rate video encoding system 400 that uses the video encoder 100 (Figure 1) with pre-calculated motion information for encoding video of a multimedia program into a set of bit streams 410-412 at various bit rates for use in the multi bit rate video streaming system 300 (Figure 3). As illustrated in the diagram, the
30 video encoding system 400 produces the set of bit streams starting from an initial highest bit rate stream 410, and including bit streams 411-412 at a succession of decreasing bit rates. For purposes of illustration, Figure 4 shows the multi bit rate video encoding system 400 producing three successively decreasing bit rate streams. However, it should be

apparent that the multi bit rate video encoding system can be designed to produce any number of streams at various bit rates.

[043] Figure 5 illustrates a process flow 500 performed by the multi bit rate video encoding system 400 for encoding the set of bit streams 410-412. At a first action, the multi bit rate video encoding system 400 first encodes an initial highest bit rate stream 410. For the initial highest bit rate stream 410, the video encoding system 400 inputs the raw video content to the video input 110 of the video encoder 100. Because this is the initial encoding of the video content, pre-calculated motion information generally is not available, and the motion information input 120 of the video encoder is zero. In this case, the motion estimator 258 (Figure 2) of the video encoder 100 performs the motion estimation process to produce motion information for motion compensation of the inter-frames of the video content for the highest bit rate stream. Alternatively, in the case that pre-calculated motion information is available (such as where the motion information is pre-calculated using a multi core motion pre-calculation discussed below), the pre-calculated motion information is provided to the input 120 of the video encoder 100. The video encoder can then choose (e.g., via motion information selection switch 256 in Figure 2) to use the input motion information for motion compensation of the inter-frames during encoding.

[044] In a repeated loop of actions 511-518, the multi-rate video encoding system next encodes each subsequent bit rate stream 411-412. For these subsequent streams 411-412 at successively decreasing bit rates, the multi bit rate video encoding system 400 again provides the same raw video content at input 110 of the video encoder 100. In addition, the motion information used for motion compensation of each stream 410-412 is provided at the output 140 of the video encoder 100, and passed on to the motion information input 120 of the video encoder of the next stream 411-412 as shown at action 512. Thus, for example, the motion information calculated by the motion estimator 258 in the video encoder for the initial stream 410 is passed on for use by the video encoder of the next lower bit rate stream 411.

[045] As indicated by actions 513-517, the video encoder for each subsequent bit rate stream then chooses (e.g., via motion information selection switch 256) whether to use the motion information passed on from the previous stream to its input 120 (action 515), or to perform a new motion estimation (action 516). In general, the decision whether to use the pre-calculated motion information or perform motion estimation anew is made based on whether the encoding parameters are being changed to achieve the next lower bit rate to an

extent that the motion information may not be sufficiently reliable to perform motion compensation with a desired encoding efficiency and quality. In one example implementation of the multi bit rate video encoding system 400, the motion information selection can be based on whether the video resolution is changed to achieve the new lower bit rate for the current stream 411-412. Accordingly, the video encoder determines the video resolution to be used for encoding the stream at action 513. At decision 514, the motion information selection switch 256 decides which motion information to use based on whether the video resolution is resized relative to that of the previous stream. If the video resolution remains unchanged, the video encoder chooses to encode the subsequent stream using the motion information passed on from encoding the previous stream as shown by action 515. However, if the video resolution is changed, the video encoder instead uses its motion estimator 258 to perform a new motion estimation (action 516) for the video content with the reference frames being intra-coded at the new video resolution. At action 517, the video encoder then encodes the stream using this new motion information. The newly calculated motion information would then be passed to the video encoder of the next lower bit rate stream (at action 512 in a next iteration of the loop 511-518 for encoding the next subsequent stream).

[046] In this way, the motion information calculated from motion estimation during encoding of a previous bit rate stream is passed on, and may be used by the multi bit rate video encoding system 400 for encoding of inter frames of a subsequent lower bit rate stream.

III. Multi Core Motion Calculation Video Encoder

[047] Similar to the motion estimator 258 of the video encoder 100 discussed above, motion estimation is typically performed for blocks or like other set of samples in an inter-frame picture with respect to a reference picture that generally is reconstructed (i.e., as would be done by a video decoder) from a previously encoded frame (e.g., typically a previously encoded intra-frame of the video). The motion estimation typically involves a search for a most closely matching block in the reference picture. In other words, the motion estimation attempts to locate a block in the reference picture that produces a smallest residual for a given block of the inter frame picture, and yields motion information (such as a motion vector) that is the spatial displacement between the given inter frame picture block and the closest matching block of the reference frame. The reason for the motion estimation to use reference frames that are reconstructed intra coded

frames is that only such previously encoded frames will be identically available in the decoder for use in reconstructing inter-coded frames.

[048] The following section presents a technique that uses a multi core processor computer to perform motion estimation pre-calculation for raw video content that can then be used as pre-calculated motion information input to the video encoder 100. Unlike conventional motion estimation, this motion estimation pre-calculation is performed prior to any encoding (including intra-frame coding) of the video content. Accordingly, there are no previously encoded frames yet available that could be reconstructed for use as the reference frames. This multi core motion information pre-calculation technique uses frames of the raw video content for the reference pictures of the motion estimation, instead of reconstructed intra-coded frame pictures. Due to quantization and other compression loss when later encoded by the video encoder 100, these raw video frames used as reference frames generally cannot be identically reconstructed at the decoder. Because of the use of raw video frames as reference pictures in this multi core motion estimation pre-calculation, the motion information may not be as accurate as motion estimation during video encoding based on reconstructed, previously encoded intra frames, which may result in less compression efficiency when used for inter frame encoding.

[049] Figure 6 illustrates one example implementation of a video encoding system 600 that includes a motion information pre-calculation on a multi core processor computer, together with encoding using the pre-calculated motion information input video encoder 100 (Figure 1). The video encoding system 600 performs motion estimation as a pre-calculation (prior to video encoding).

[050] The video encoding system 600 achieves a performance advantage by dividing the computation load of the motion estimation among the separate CPU cores 620-623 of a multi core processor computer 610. The multi core processor computer 610 can have any number “n” of CPU cores on which the motion estimation is processed. For example, presently available multi core computers manufactured by Intel and other have 4, 8, 16, 32, 64 and even many more processor cores.

[051] For the motion estimation, the CPU cores 620-623 are assigned to each perform the motion estimation of a separate inter-frame out of a group of frames of the raw video content. In this way, the motion estimation pre-calculation avoids memory contention between the CPU cores. Separate inter-frames of the raw video content (labeled F0 ... Fn in the drawing) are provided along with the respective reference frame to separate CPU cores of the computer 610. The CPU cores perform motion estimation on their respective

inter frame picture and produce motion information (such as a set of motion vectors for the blocks of the inter frame). This is then repeated for a next group of inter frame pictures of the raw video content (e.g., in groups of “n” frames per iteration) to produce motion information for the full video sequence or program.

5 [052] The motion information produced by the motion estimation pre-calculation using the CPU cores can then be input to the video encoder 100 (Figure 1) along with the raw video content. The video encoder 100 then proceeds to encode the video, including encoding the inter frames of the video content using the motion information from the multi core motion estimation pre-calculation. Because the video encoder 100 uses motion
10 information from the multi-core motion estimation pre-calculation, the video encoder 100 avoids performing the motion estimation (in motion estimator 258) itself, which saves a substantial part of the video encoding. Because the motion estimation is performed essentially in parallel among the multi CPU cores as a pre-computation rather than sequentially as the video frames are encoded, this multi core video encoder achieves much
15 faster encoding of the video.

IV. Representative Computing Environment

[053] Figure 7 illustrates a generalized example of a suitable computing environment 700 in which described embodiments, techniques, and technologies may be implemented. The computing environment 700 is not intended to suggest any limitation as to scope of use or
20 functionality of the technology, as the technology may be implemented in diverse general-purpose or special-purpose computing environments. For example, the disclosed technology may be implemented with other computer system configurations, including hand held devices, multiprocessor systems, microprocessor-based or programmable consumer electronics, network PCs, minicomputers, mainframe computers, and the like.
25 The disclosed technology may also be practiced in distributed computing environments where tasks are performed by remote processing devices that are linked through a communications network. In a distributed computing environment, program modules may be located in both local and remote memory storage devices.

[054] With reference to Figure 7, the computing environment 700 includes at least one
30 central processing unit 710 and memory 720. For the multi core motion information precalculation discussed above, the computer includes a multi core CPU having plural CPU cores. In Figure 7, this most basic configuration 730 is included within a dashed line. The central processing unit 710 executes computer-executable instructions and may be a real or a virtual processor. In a multi-processing system, multiple processing units

execute computer-executable instructions to increase processing power and as such, multiple processors can be running simultaneously. The memory 720 may be volatile memory (e.g., registers, cache, RAM), non-volatile memory (e.g., ROM, EEPROM, flash memory, etc.), or some combination of the two. The memory 720 stores software 780 that
5 can, for example, implement the technologies described herein. A computing environment may have additional features. For example, the computing environment 700 includes storage 740, one or more input devices 750, one or more output devices 760, and one or more communication connections 770. An interconnection mechanism (not shown) such as a bus, a controller, or a network, interconnects the components of the computing
10 environment 700. Typically, operating system software (not shown) provides an operating environment for other software executing in the computing environment 700, and coordinates activities of the components of the computing environment 700.

[055] The storage 740 may be removable or non-removable, and includes magnetic disks, magnetic tapes or cassettes, CD-ROMs, CD-RWs, DVDs, or any other medium
15 which can be used to store information and which can be accessed within the computing environment 700. The storage 740 stores instructions for the software 780, which can implement technologies described herein.

[056] The input device(s) 750 may be a touch input device, such as a keyboard, keypad, mouse, pen, or trackball, a voice input device, a scanning device, or another device, that
20 provides input to the computing environment 700. For audio, the input device(s) 750 may be a sound card or similar device that accepts audio input in analog or digital form, or a CD-ROM reader that provides audio samples to the computing environment 700. The output device(s) 760 may be a display, printer, speaker, CD-writer, or another device that provides output from the computing environment 700.

25 [057] The communication connection(s) 770 enable communication over a communication medium (e.g., a connecting network) to another computing entity. The communication medium conveys information such as computer-executable instructions, compressed graphics information, or other data in a modulated data signal.

[058] Computer-readable media are any available media that can be accessed within a
30 computing environment 700. By way of example, and not limitation, with the computing environment 700, computer-readable media include memory 720, storage 740, the communication medium, and combinations of any of the above. As should be readily understood, the term computer-readable storage media includes the media for data storage

51017-28

such as memory 720 and storage 740, and not simply transmission media such as modulated data signals.

[059] Any of the methods described herein can be performed via one or more computer-readable media (e.g., storage or other tangible media) comprising (e.g., having or storing) 5 computer-executable instructions for performing (e.g., causing a computing device, audio and/or video processing device, or computer to perform) such methods. Operation can be fully automatic, semi-automatic, or involve manual intervention.

[060] Having described and illustrated the principles of our innovations in the detailed description and accompanying drawings, it will be recognized that the various 10 embodiments can be modified in arrangement and detail without departing from such principles. It should be understood that the programs, processes, or methods described herein are not related or limited to any particular type of computing environment, unless indicated otherwise. Various types of general purpose or specialized computing environments may be used with or perform operations in accordance with the teachings 15 described herein. Elements of embodiments shown in software may be implemented in hardware and vice versa.

[061] In view of the many possible embodiments to which the principles of our invention may be applied, we claim as our invention all such embodiments as may come within the scope of the following claims and equivalents thereto.

We claim:

1. A method of processing raw video content input for encoding into multi bit rate video, the method comprising:

receiving raw video content (110) for encoding as compressed video bit streams

5 (410, 411, 412) at a succession of progressively lower bit rates;

encoding (510) an initial compressed video bit stream (410) at a highest bit rate, said encoding the initial compressed video bit stream producing (512) motion information for encoding of inter frame pictures of the video;

10 for encoding subsequent compressed video bit streams at the progressively lower bit rates,

passing the motion information produced from encoding a preceding compressed video bit stream to the encoding of the subsequent compressed video bit stream;

15 determining (514) whether a condition based on the video encoding parameters for the subsequent compressed video bit stream is met;

if said condition is met, encoding (515) the subsequent compressed video bit stream (411, 412) wherein inter frames of the subsequent compressed video bit stream are encoded using the motion information passed from encoding the preceding compressed video bit stream; and

20 if said condition is not met, producing new motion information by performing motion estimation (516) based on reconstructed reference frames of the subsequent compressed video bit stream, and encoding (517) the subsequent compressed video bit stream wherein inter frames of the subsequent compressed video bit stream are encoded using the new motion information; and

25 producing an output multi bit rate video comprising the initial and subsequent compressed video bit streams.

2. The method of claim 1 wherein said condition is that the video resolution remain unchanged between the preceding compressed video bit stream and the subsequent compressed video bit stream.

30 3. A method of encoding video on a computer having a multi core CPU, the method comprising:

producing motion information for groups of inter frame pictures of a video as a motion estimation pre-calculation prior to encoding of the video, wherein said producing motion information comprises:

for each inter frame picture in a group, performing motion estimation on a separate one of plural CPU cores (620-623) of the computer, said motion estimation producing motion information for plural sets of elements of the respective inter frame picture;

5 encoding the video, wherein said encoding comprises encoding the inter frame pictures based on the motion information produced by the motion estimation pre-calculation; and

producing an output compressed video bit stream (130) containing the encoded video.

10 4. The method of claim 3 further comprising:

encoding the video as one or more additional streams at progressively lower bit rates, wherein said encoding the additional streams comprises:

15 passing the motion information produced from encoding a preceding compressed video bit stream to the encoding of a subsequent compressed video bit stream;

determining whether a condition based on the video encoding parameters for the subsequent compressed video bit stream is met;

20 if said condition is met, encoding the subsequent compressed video bit stream wherein inter frames of the subsequent compressed video bit stream are encoded using the motion information passed from encoding the preceding compressed video bit stream; and

25 if said condition is not met, producing new motion information by performing motion estimation based on reconstructed reference frames of the subsequent compressed video bit stream, and encoding the subsequent compressed video bit stream wherein inter frames of the subsequent compressed video bit stream are encoded using the new motion information; and
producing an output multi bit rate video comprising the initial and subsequent compressed video bit streams.

30 5. The method of claim 4 wherein said condition is that the video resolution remain unchanged between the preceding compressed video bit stream and the subsequent compressed video bit stream.

6. A video processing system for encoding compressed video streams for multiple bit rate video streaming, the system comprising:

a memory (720) for storing raw video content to be encoded;

a computer processing means (710) programmed for encoding the raw video content as a plurality of compressed video streams, wherein said encoding the compressed video streams comprises:

5 producing original motion information for encoding inter-frame pictures of the raw video content via motion estimation;

 encoding a top quality compressed video bit stream (410) from the raw video content using the original motion information for encoding inter-frame pictures of the raw video content;

10 deciding (513) a video resolution for encoding a subsequent compressed video bit stream (411, 412) at a lower bit rate;

 if the video resolution for encoding the subsequent compressed video bit stream is unchanged, encoding (515) the subsequent compressed video bit stream from the raw video content again using the original motion information for encoding inter-frame pictures of the raw video content; and

15 otherwise if the video resolution for encoding the subsequent compressed video bit stream is resized from that of the top quality compressed video bit stream, then producing (516) new motion information for encoding inter-frame pictures of the raw video content via motion estimation relative to intra-frame reference pictures of the subsequent compressed video bit stream, and encoding (517) the
20 subsequent compressed video bit streams from the raw video content using the new motion information for encoding inter-frame pictures of the raw video content.

7. The video processing system of claim 6 wherein the computer processing means is a multi core computer processor having multiple processing cores, and wherein said producing original motion information comprises, as a pre-calculation prior to
25 encoding intra-frame pictures of the raw video content, performing motion estimation for each inter frame picture in a group of inter frame pictures of the raw video content on a separate one of the processing cores.

8. The video processing system of claim 7 said performing motion estimation for a group of inter frame pictures of the raw video content comprises:

30 assigning each inter frame picture of the group to a separate processing core of the multi core computer processor; and

 performing motion estimation for each respective inter frame picture on its assigned separate processing core.

51331-1055

9. A computer readable medium having computer executable instructions stored thereon for execution by one or more computers, that when executed implement the method according to claim 1 or claim 2.

10. A computer readable medium having computer executable
5 instructions stored thereon for execution by one or more computers, that when executed implement the method according to claim 3 or claim 4.

1 / 7

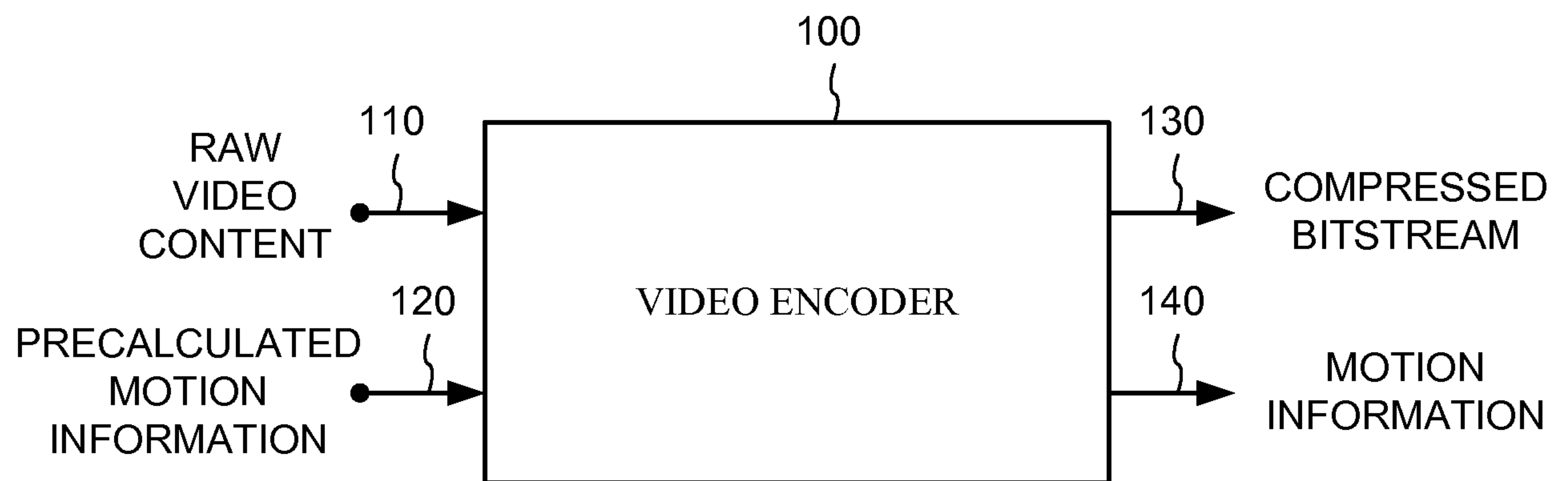


FIG. 1

2 / 7

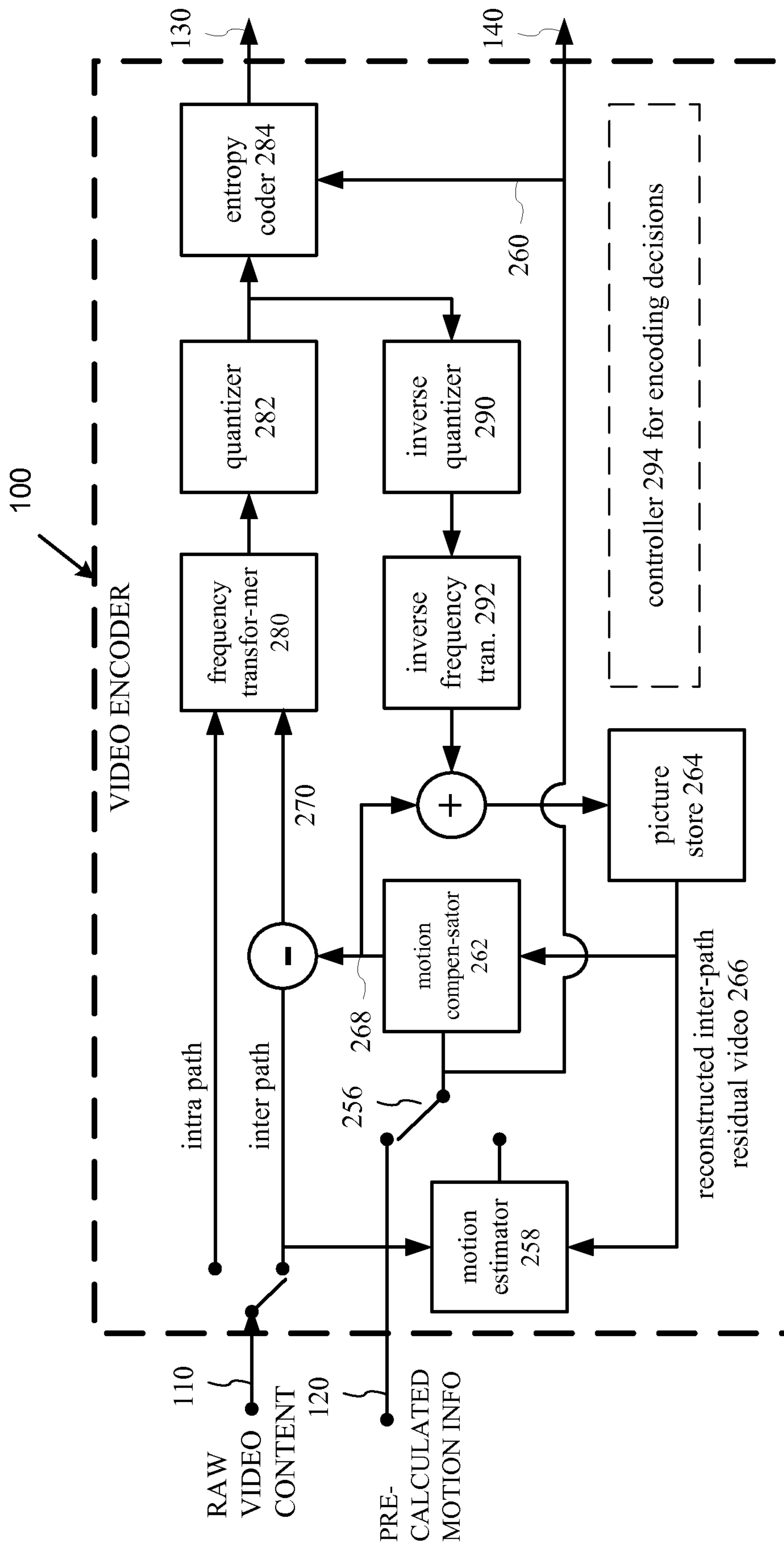
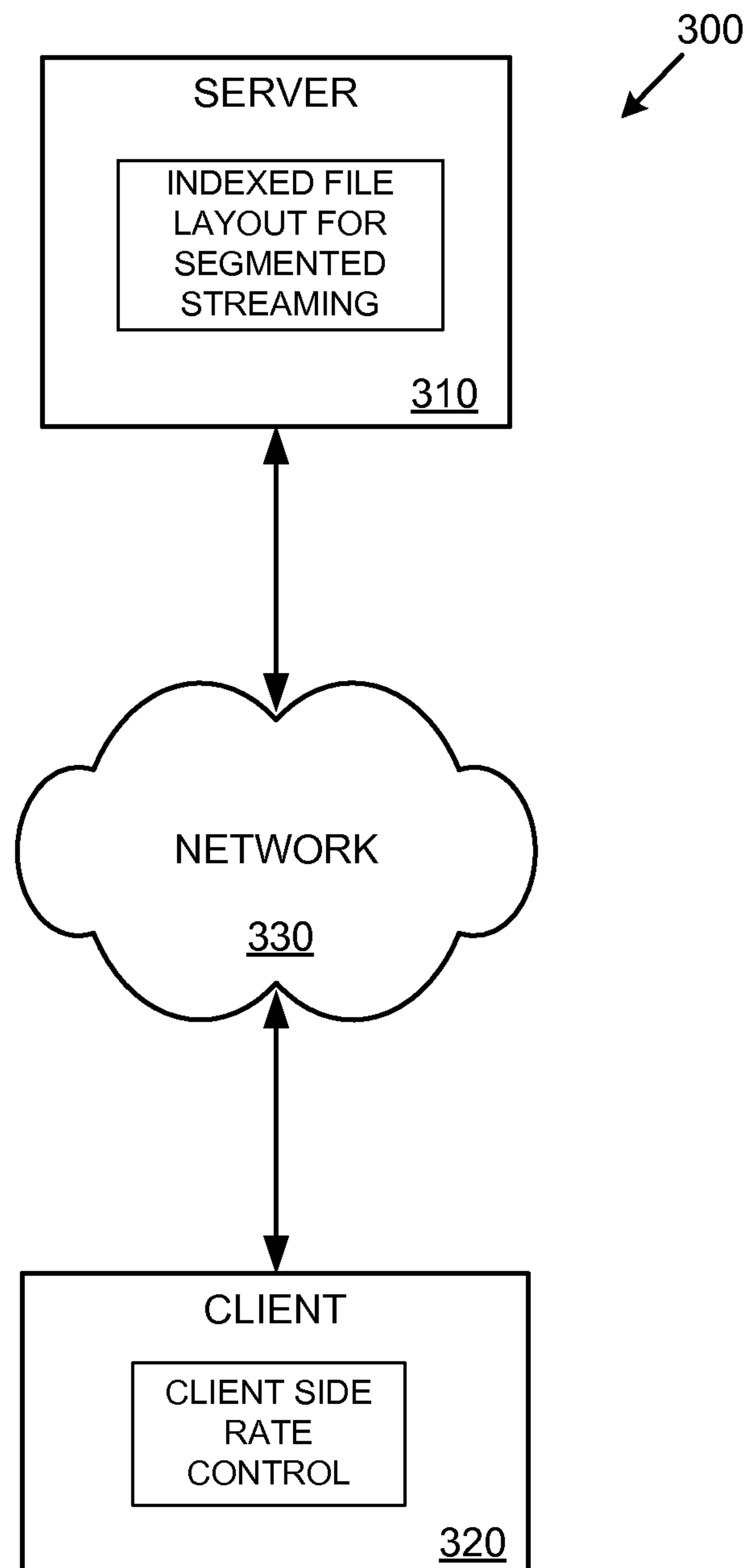


FIG. 2

3 / 7

**FIG. 3**

4 / 7

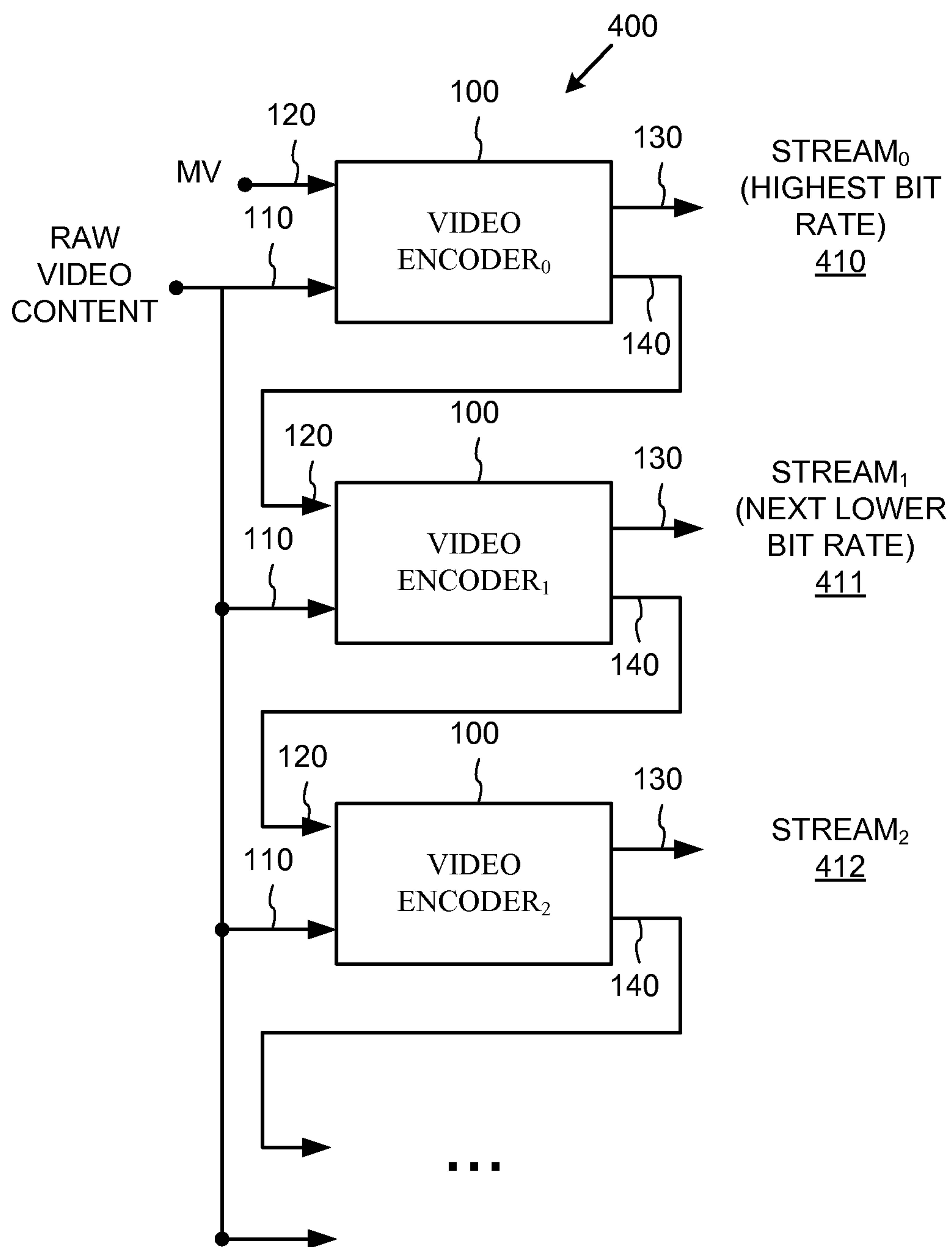


FIG. 4

5 / 7

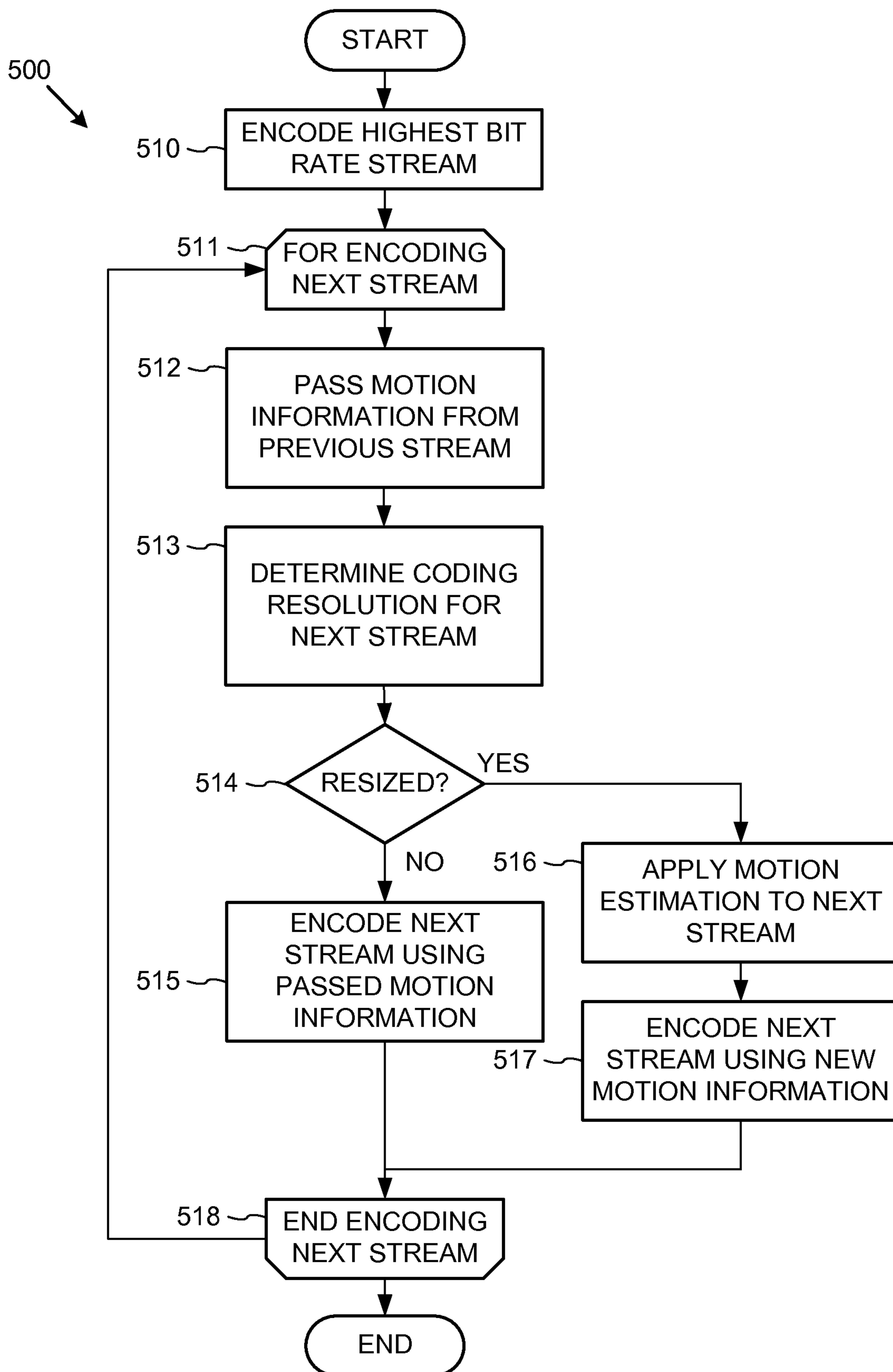


FIG. 5

6 / 7

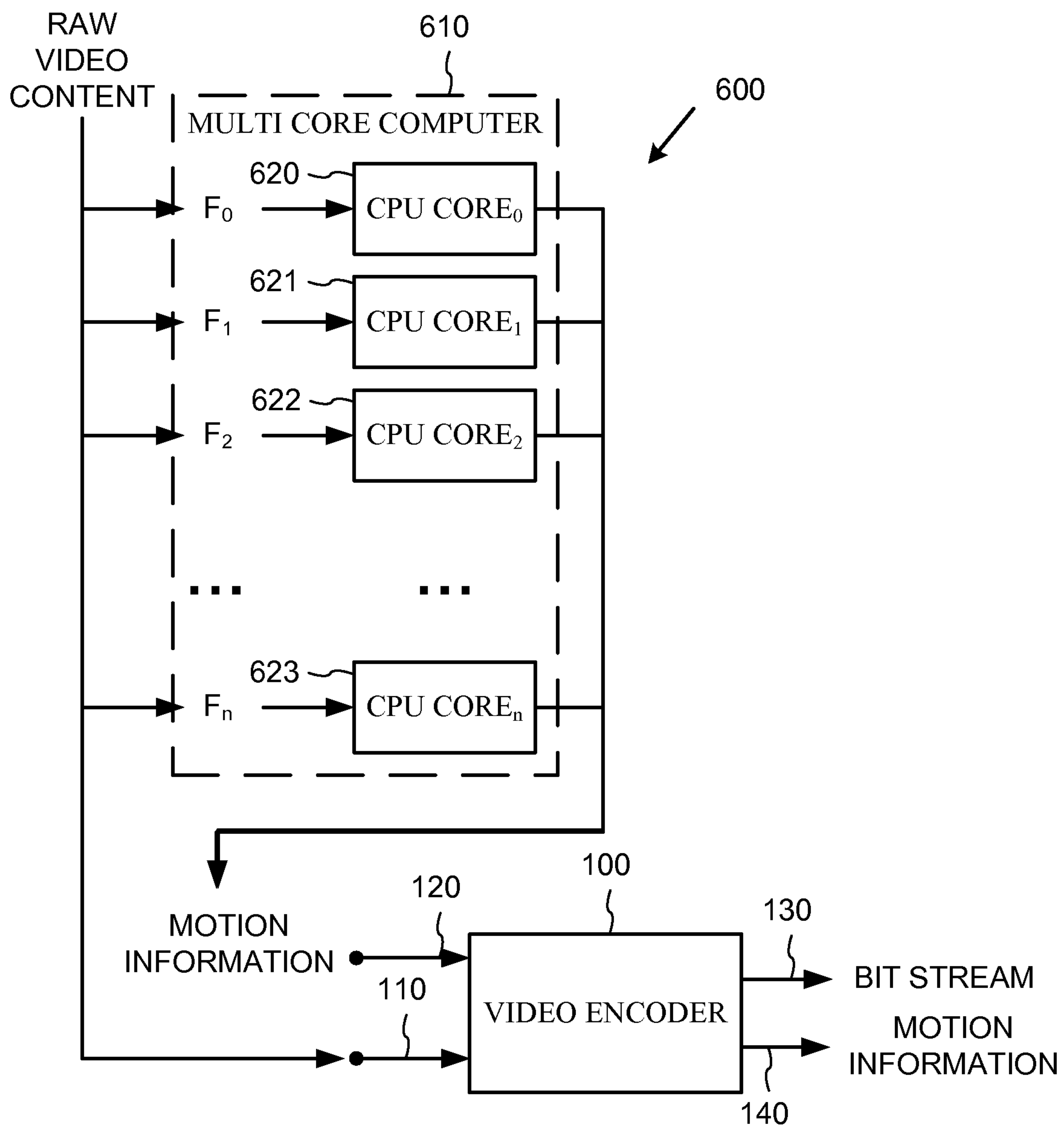


FIG. 6

7 / 7

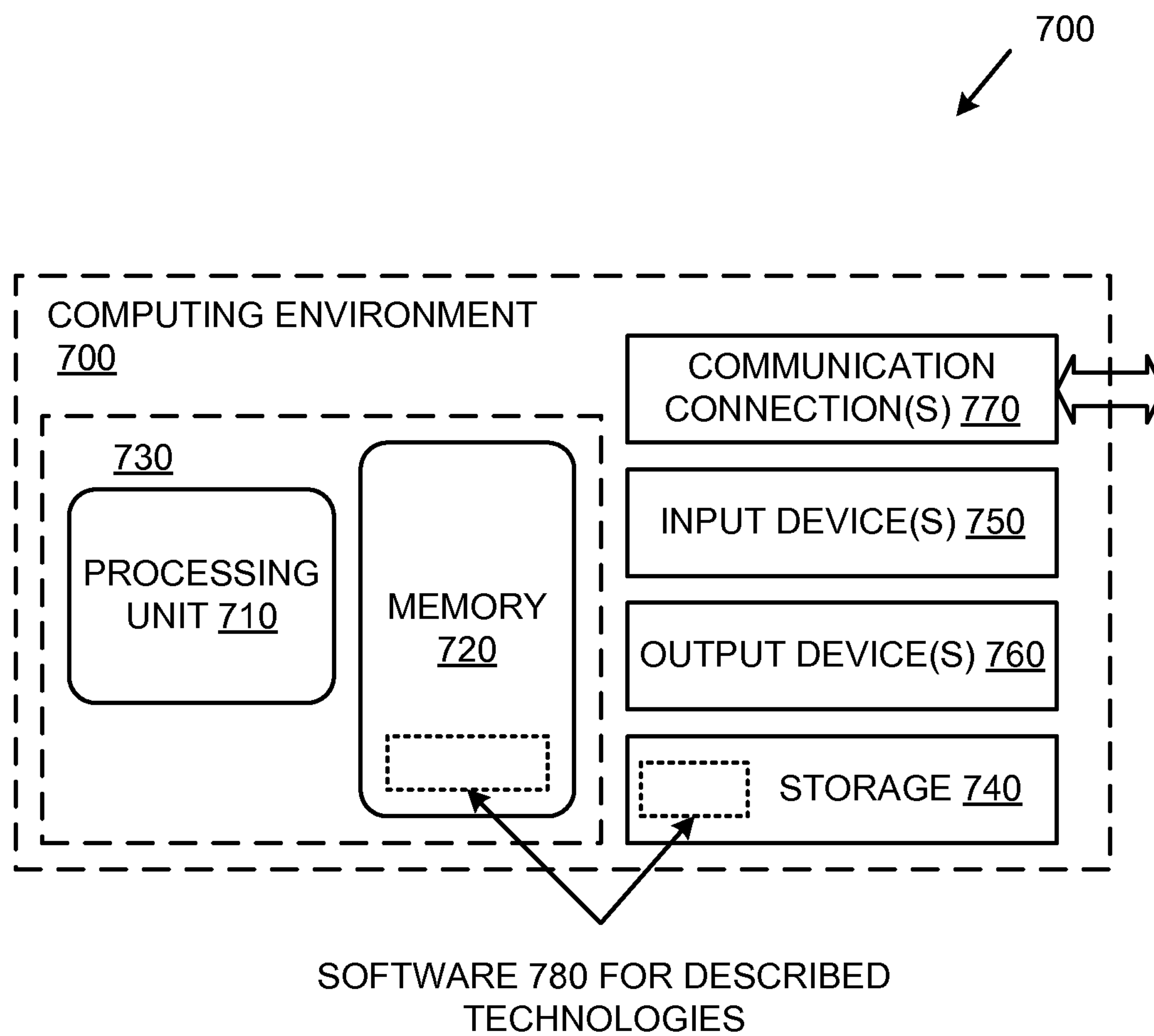


FIG. 7

