

**(12) STANDARD PATENT**  
**(19) AUSTRALIAN PATENT OFFICE**

**(11) Application No. AU 2016243988 B2**

(54) Title  
**Techniques for optimizing bitrates and resolutions during encoding**

(51) International Patent Classification(s)  
**H04N 21/234** (2011.01)      **H04N 21/2343** (2011.01)

(21) Application No: **2016243988**      (22) Date of Filing: **2016.03.10**

(87) WIPO No: **WO16/160295**

(30) Priority Data

(31) Number **14/673,621**      (32) Date **2015.03.30**      (33) Country **US**

(43) Publication Date: **2016.10.06**  
(44) Accepted Journal Date: **2019.02.07**

(71) Applicant(s)  
**Netflix, Inc.**

(72) Inventor(s)  
**Aaron, Anne;Ronca, David;Katsavounidis, Ioannis;Schuler, Andy**

(74) Agent / Attorney  
**FPA Patent Attorneys Pty Ltd, Level 43 101 Collins Street, Melbourne, VIC, 3000, AU**

(56) Related Art  
**WO 2012078965 A1**  
**SUN, H. et al., "Comparing Subjective Perceived Quality with Objective Video Quality by Content Characteristics and Bit Rates", IEEE International Conference on New Trends in Information and Service Science, 30 June-2 July 2009, Beijing**  
**WO 2012078965 A1**  
**US 20090310673 A1**  
**SUN, H. et al., "Comparing Subjective Perceived Quality with Objective Video Quality by Content Characteristics and Bit Rates", IEEE International Conference on New Trends in Information and Service Science, 30 June-2 July 2009, Beijing**  
**WO 2015183910 A1**  
**US 20090310673 A1**  
**US 8767825 B1**  
**US 8767825 B1**  
**US 8743954 B1**  
**US 8743954 B1**  
**WO 2016130318 A1**  
**WO 2016130318 A1**  
**WO 2015183910 A1**

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

(19) World Intellectual Property Organization

International Bureau



(10) International Publication Number

WO 2016/160295 A1

(43) International Publication Date

6 October 2016 (06.10.2016)

WIPO | PCT

(51) International Patent Classification:

H04N 21/234 (2011.01) H04N 21/2343 (2011.01)

(21) International Application Number:

PCT/US2016/021649

(22) International Filing Date:

10 March 2016 (10.03.2016)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

14/673,621 30 March 2015 (30.03.2015) US

(71) Applicant: NETFLIX, INC. [US/US]; 100 Winchester Circle, Los Gatos, California 95032 (US).

(72) Inventors: AARON, Anne; 256 Marmona Drive, Menlo Park, California 94025 (US). RONCA, David; 1459 McBain Ave., Campbell, California 95008 (US). KATSA-VOUNDIS, Ioannis; 10754 Westminster Ave., Los Angeles, California 90034 (US). SCHULER, Andy; 943 Goodwin Ave, San Jose, California 95128 (US).

(74) Agents: CAREY, John C. et al.; Artegis Law Group, LLP, 7710 Cherry Park Drive, Suite T104, Houston, Texas 77095 (US).

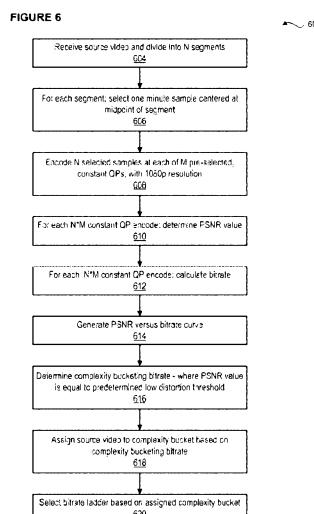
(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, 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, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, 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, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— with international search report (Art. 21(3))

(54) Title: TECHNIQUES FOR OPTIMIZING BITRATES AND RESOLUTIONS DURING ENCODING



(57) Abstract: In one embodiment of the present invention, an encoding bitrate ladder selector tailors bitrate ladders to the complexity of source data. Upon receiving source data, a complexity analyzer configures an encoder to repeatedly encode the source data-set using a constant quantization parameter to a different value for each encode. The complexity analyzer processes the encoding results to determine an equation that relates a visual quality metric to an encoding bitrate. The bucketing unit solves this equation to estimate a bucketing bitrate at a predetermined value of the visual quality metric. Based on the bucketing bitrate, the bucketing unit assigns the source data to a complexity bucket having an associated, predetermined bitrate ladder. Advantageously, sagaciously selecting the bitrate ladder enables encoding that optimally reflects tradeoffs between quality and resources (e.g., storage and bandwidth) across a variety of source data types instead of a single, "typical" source data type.

WO 2016/160295 A1

## TECHNIQUES FOR OPTIMIZING BITRATES AND RESOLUTIONS DURING ENCODING

### CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims benefit of United States patent application serial 5 number 14/673,621, filed March 30, 2015, which is hereby incorporated herein by reference.

### BACKGROUND OF THE INVENTION

#### **Field of the Invention**

[0002] Embodiments of the present invention relate generally to computer science 10 and, more specifically, to techniques for optimizing bitrates and resolutions during encoding.

#### **Description of the Related Art**

[0003] Efficiently and accurately encoding source data is essential for real-time 15 delivery of video content. In operation, after the encoded data is received at an endpoint machine, the encoded data is decoded and viewed or otherwise further processed. To increase compression rates and/or reduce the size of the encoded data, many encoding processes leverage lossy data compression techniques that eliminate selected information, typically enabling only approximate reconstruction of the source data. Notably, as the encoder eliminates information, the resolution of the 20 encoded data decreases and, consequently, the likelihood that the approximate reconstruction has the visual quality viewers expect and desire also decreases.

[0004] In operation, encoders are often configured to implement a fixed bitrate 25 ladder that makes tradeoffs between resources consumed during the encoding/decoding process (e.g., processing time, bandwidth, storage, etc.) and visual quality. Each “rung” in the bitrate ladder represents a different bitrate and resolution. In general, given an available bitrate, the encoder selects the encoding bitrate and resolution based on the bitrate ladder, and then generates encoding data at the determined bitrate and resolution.

2016243988  
17 Dec 2018

[0005] In practice, a bitrate ladder is tuned to generate encoded data having the requisite level of quality for “typical” source data. However, in situations where source data differs noticeably from “typical” source data, the tradeoffs represented by the bitrate ladder may not be appropriate. For example, if the bitrate ladder is designed

5 to optimize tradeoffs for simple cartoons, and the source data is a detailed action movie, then the tradeoffs that the bitrate ladder imposes during encoding may result in unacceptably poor visual quality. Conversely, if the bitrate ladder is designed to optimize tradeoffs for detailed action movies, and the source data is a simple cartoon, then the tradeoffs that the bitrate ladder imposes during encoding may dramatically 10 increase resource burdens, such as storage and bandwidth usage, without noticeably increasing visual quality.

[0006] As the foregoing illustrates, what is needed in the art are more effective techniques for selecting bitrates and resolutions when encoding source data.

[0006A] Reference to any prior art in the specification is not an acknowledgment or 15 suggestion that this prior art forms part of the common general knowledge in any jurisdiction or that this prior art could reasonably be expected to be understood, regarded as relevant, and/or combined with other pieces of prior art by a skilled person in the art.

## SUMMARY OF THE INVENTION

20 [0007] In a first aspect of the present invention there is provided a computer-implemented method for selecting a bitrate ladder for encoding source data, the method comprising: selecting a set of parameter values for a quantization parameter; for each parameter value, performing an encoding process that configures an encoder to encode source data to 25 generate a set of encoded data while maintaining the quantization parameter at the parameter value; for each set of encoded data, determining a value of a video quality metric and a corresponding bitrate; deriving a relationship between the video quality metrics and the corresponding bitrates; determining based on the relationship, a bucketing bitrate corresponding to a first video quality metric, wherein the first video 30 quality metric is equal to a predetermined threshold; bit rate at which the video quality

17 Dec 2018

2016243988

metric is equal to a predetermined threshold based on the relationship; and determining that the bucketing bitrate falls within a bitrate range associated with a first bitrate ladder included in a plurality of bitrate ladders, wherein each bitrate ladder included in the plurality of bitrate ladders is associated with a different bitrate range and includes a different plurality of bitrate-resolution pairs; and selecting a the first bitrate ladder for encoding the source data based on the bucketing bitrate.

**[0007A]** In a second aspect of the present invention there is provided a non-transitory computer-readable storage medium including instructions that, when executed by a processing unit, cause the processing unit to select a bitrate ladder for

encoding source data by performing the steps of: selecting a set of parameter values for a quantization parameter; for each parameter value, performing an encoding process that configures an encoder to encode source data to generate a set of encoded data while maintaining the quantization parameter at the parameter value; for each set of encoded data, determining a value of a video quality metric and a corresponding bitrate; deriving a relationship between the video quality metrics and the corresponding bitrates; determining, based on the relationship, a bucketing bitrate corresponding to a first video quality metric, wherein the first video quality metric is equal to a predetermined threshold; determining that the bucketing bitrate falls within a bitrate range associated with a first bitrate ladder included in a plurality of bitrate ladders, wherein each bitrate ladder included in the plurality of bitrate ladders is associated with a different bitrate range and includes a different plurality of bitrate-resolution pairs; and selecting the first bitrate ladder for encoding the source data.

**[0007B]** In a third aspect of the present invention there is provided a system configured select a bitrate ladder for encoding source data, the system comprising: a

memory storing instructions; a processor executing instructions to: select a set of parameter values for a quantization parameter; for each parameter value, perform an encoding process that configures an encoder to encode source data to generate a set of encoded data while maintaining the quantization parameter at the parameter value; for each set of encoded data, determine a value of a video quality metric and a corresponding bitrate; derive a relationship between the video quality metrics and the corresponding bitrates; determine, based on the relationship, a bucketing bitrate

17 Dec 2018

2016243988

range corresponding to a first video quality metric, wherein the first video quality metric is equal to a predetermined threshold; determine that the bucketing bitrate falls within a bitrate range associated with a first bitrate ladder included in a plurality of bitrate ladders, wherein each bitrate ladder included in the plurality of bitrate ladders is associated with a different bitrate range and includes a different plurality of bitrate-resolution pairs; and select the first bitrate ladder for encoding the source data.

[0008] One advantage of the disclosed bitrate ladder selection techniques is that

these techniques enable the selection of complexity-tuned bitrates and resolutions when encoding source data. Notably, because the disclosed techniques tailor the

bitrate ladder to reflect the complexity of the source data, tradeoffs between encoding quality and encoding resources are optimized compared to conventional techniques

that implement a constant bitrate ladder irrespective of the complexity of the source data.

### **BRIEF DESCRIPTION OF THE DRAWINGS**

**[0009]** So that the manner in which the above recited features of the present invention

5 can be understood in detail, a more particular description of the invention, briefly summarized above, may be had by reference to embodiments, some of which are illustrated in the appended drawings. It is to be noted, however, that the appended drawings illustrate only typical embodiments of this invention and are therefore not to be considered limiting of its scope, for the invention may admit to other equally  
10 effective embodiments.

**[0010]** Figure 1 is a conceptual illustration of a system configured to implement one or more aspects of the present invention;

**[0011]** Figure 2 is a block diagram illustrating the encoding bitrate ladder selector of Figure 1, according to one embodiment of the present invention;

15 **[0012]** Figure 3 is a conceptual illustration of the complexity buckets and the bitrate ladders implemented by the encoding bitrate ladder selector of Figure 2, according to one embodiment of the present invention;

**[0013]** Figure 4 is a conceptual illustration of a peak signal-to-noise ratio (PSNR) curve for a simple source constructed by the complexity analyzer of Figure 2,  
20 according to one embodiment of the present invention;

**[0014]** Figure 5 is a conceptual illustration of a peak signal-to-noise ratio (PSNR) curve for a complex source constructed by the complexity analyzer of Figure 2, according to one embodiment of the present invention; and

25 **[0015]** Figure 6 is a flow diagram of method steps for selecting and implementing a bitrate ladder while encoding video source data, according to one embodiment of the present invention.

## **DETAILED DESCRIPTION**

[0016] In the following description, numerous specific details are set forth to provide a more thorough understanding of the present invention. However, it will be apparent to one of skilled in the art that the present invention may be practiced without one or 5 more of these specific details.

### **System Overview**

[0017] Figure 1 is a conceptual illustration of a system 100 configured to implement one or more aspects of the present invention. As shown, the system 100 includes a virtual private cloud (i.e., encapsulated shared resources, software, data, etc.) 102 10 connected to a variety of devices capable of transmitting input data and/or displaying video. Such devices include, without limitation, a desktop computer 102, a smartphone 104, and a laptop 106. In alternate embodiments, the system 100 may include any number and/or type of input, output, and/or input/output devices in any combination.

15 [0018] The virtual private cloud (VPC) 102 includes, without limitation, any number and type of compute instances 110. The VPC 102 receives input user information from an input device (e.g., the laptop 106), one or more computer instances 110 operate on the user information, and the VPC 102 transmits processed information to the user. The VPC 102 conveys output information to the user via display 20 capabilities of any number of devices, such as a conventional cathode ray tube, liquid crystal display, light-emitting diode, or the like.

[0019] In alternate embodiments, the VPC 102 may be replaced with any type of 25 cloud computing environment, such as a public or a hybrid cloud. In other embodiments, the system 100 may include any distributed computer system instead of the VPC 102. In yet other embodiments, the system 100 does not include the VPC 102 and, instead, the system 100 includes a single processing or multi-processing unit.

[0020] As shown for the compute instance 110<sub>0</sub>, each compute instance 110 includes 30 a central processing unit (CPU) 112, a graphics processing unit (GPU) 114, and a memory 116. In operation, the CPU 112 is the master processor of the compute

instance 110, controlling and coordinating operations of other components included in the compute instance 110. In particular, the CPU 112 issues commands that control the operation of the GPU 114. The GPU 114 incorporates circuitry optimized for graphics and video processing, including, for example, video output circuitry. In 5 various embodiments, GPU 114 may be integrated with one or more of other elements of the compute instance 110. The memory 116 stores content, such as software applications and data, for use by the CPU 112 and the GPU 114 of the compute instance 110.

[0021] In general, the compute instances 110 included in the VPC 102 are 10 configured to implement one or more applications. More specifically, the compute instances 110 included in the VPC 102 are configured to encode source data 105, such as a video file. As shown, compute instance 110<sub>0</sub> is configured as a source inspector 110, and compute instances 110<sub>1</sub>-110<sub>N</sub> are configured as an encoder 140. In alternate embodiments, source inspector 110 may include more compute instances 15 110, and encoder 140 may include only a single compute instance 110.

[0022] The source inspector 110 receives the source data 105 and performs any 20 number of pre-encoding operations, including configuring the encoder 140. For example, in some embodiments, the encoder 140 is a parallel chunk encoder. In such embodiments the source inspector 110 breaks the source into multiple source 25 chunks prior to routing the source chunks to compute instances 140 included in the parallel chunk encoder.

In general, the encoder 140 includes multiple modes and settings that enable 30 customization of the encoding operations (e.g., compression algorithms). In particular, the encoder 140 implements, without limitation, both a constant bitrate mode and a constant quality encoding mode. In the constant bitrate mode, the encoder 140 attempts to maintain a target bitrate throughout the encoding process. If the encoder 140 detects that the encoded bitrate is greater than the target bitrate, then the encoder 140 reduces the quality of encoded data 145. If the encoder 140 detects that the encoded bitrate is less than the target bitrate, then the encoder 140 increases the quality of the encoded data 145. By contrast, in the constant quality encoding mode, the encoder 140 attempts to maintain a constant quality throughout

the encoding process. The encoder 140 measures the quality of the encoded data 145 and varies a quantization parameter (QP) that defines the amount of signal detail to include in the encoded data 145 to maintain the encoded quality at the target quality. Alternatively, the encoder 140 sets the QP to a constant value throughout the 5 encoding process.

**[0023]** As persons skilled in the art will recognize, the configuration of the encoder 140 dramatically impacts the required resources, such as bandwidth and storage, and the quality of the encoded data 145. Configuring the encoder 140 to optimize the encoded data 145 to satisfy an acceptable perceived visual quality without wasting 10 required resource across different types of the source data 105 is difficult. Conventional approaches, such as those leveraging a constant bitrate ladder and encoding at a target bitrate, optimize the quality/resource tradeoff for some types of source data 105, but produce poor results for other types of source data 105.

**[0024]** For this reason, the source inspector 110 includes an encoding bitrate 15 ladder selector 130. In operation, the encoding bitrate ladder selector 130 leverages the encoder 140 to estimate the complexity of the source data 105 and then assigns the source data 105 to a complexity bucket 132. Subsequently, the encoding bitrate ladder selector 130 identifies a bitrate ladder 134 that is associated with the selected complexity bucket 132 as a source-tune bitrate ladder 135. In this fashion, the 20 encoding bitrate ladder selector 130 enables the quality/resource tradeoff to be optimized to reflect the complexity of the source data 105.

**[0025]** In general, the encoding bitrate ladder selector 130 may include any 25 number of complexity buckets 132, where each of the complexity buckets 132 represents a different range of complexities for the source data 105. For example, in some embodiments, the “simple” complexity bucket 132 represents the lowest range of complexities, such as source data 105 for a basic cartoon. Each of the complexity buckets 132 is associated a particular one of the bitrate ladders 134 that is tailored to optimize the encoding tradeoffs for source data 105 of the complexities represented by the complexity bucket 132.

**[0026]** The encoding bitrate ladder selector 130 may implement the complexity buckets 132 and the bitrate ladders 134 in any technically feasible fashion. In some embodiments, the encoding bitrate ladder selector 130 includes an initialization unit that establishes the complexity buckets 132 and the bitrate ladders 134 prior to 5 processing any source data, including the source data 105. In other embodiments, the initialization unit is not included in the encoding bitrate ladder selector 130, but is a unit included in the system 100. In yet other embodiments, the initialization unit is not included in system 100.

**[0027]** In one embodiment, the initialization unit provides encoding quality 10 feedback information that is manually evaluated to generate the optimized complexity buckets 132 and the bitrate ladders 134. First, the initialization unit receives a “typical” bitrate ladder. The initialization unit then analyzes a variety of test data (spanning the expected complexity range of the source data 105) across the bitrates included in the typical bitrate ladder. The resulting encoded test data spans both the 15 rungs of the ladder and the test data.

**[0028]** After generating the encoded test data, the initialization unit evaluates quality of each of the encoded test data. More specifically, the initialization unit applies the Tektronix® Picture Quality Analyzer (PQA) tool to the encoded test data. The PQA tool assigns a Picture Quality Rating (PQR) score to the encoded test data 20 that “closely correspond with subjective human visual assessment.” The initialization unit then graphs bitrate versus the PQR score quality, where quality is the PQR score. Based on the PQR graphs, the optimized complexity buckets 132 and the bit rate ladders 134 are selected to produce acceptable tradeoffs between resource and quality across the expected complexity of the source data 105. Among other things, 25 the optimized complexity buckets 132 and the bitrate ladders 134 may be created to satisfy the following objectives:

- Improve the video quality given the same bitrate by optimizing the resolution.
- Save on storage and bandwidth. For simple sources, the high bitrate streams may be unnecessary (no obvious quality gain) and, consequently, waste storage and bandwidth.

- Achieve better quality for complex sources, at the expense of a higher bitrate stream. For complexity buckets 132 corresponding to complicate test data, additional bitrates may be added to achieve acceptable quality.

5 [0029] The complexity buckets 132 and the bitrate ladders 134 may be generated based on the PQR graphs in any technically feasible fashion—such as manual evaluation. In alternate embodiments, the PQR score may be replaced with any quality measurement. In general, the complexity buckets 132 and the bitrate ladders 134 may be generated in any technically feasible fashion that evaluates test data  
10 across a variety of complexities and bitrates.

15 [0030] After the initialization unit finishes, the resulting complexity buckets 132 and the bitrate ladders 134 are incorporated into or communicated to the encoding bitrate ladder selector 130. In various embodiments, the complexity buckets 132 and the bitrate ladders 134 may be transferred into the system 100 via any communication method as known in the art.

### Estimating Source Complexity

20 [0031] Figure 2 is a block diagram illustrating the encoding bitrate ladder selector 130 of Figure 1, according to one embodiment of the present invention. The encoding bitrate ladder selector 130 receives the source data 105, estimates the complexity of the source data 105, and selects a corresponding source-tuned bitrate ladder 135 from the bitrate ladders 134. As part of identifying the bitrate ladder 134 that represents the optimal resource/quality tradeoff for the source data 105, the encoding bitrate ladder selector 130 leverages the encoder 140 and the complexity buckets 132.

25 [0032] As shown, the encoding bitrate ladder selector 130 includes, without limitation, a sample extractor 210, a complexity analyzer 220, and a bucketing unit 260. Upon receiving the source data 105, the sample extractor 210 partitions the source data 105 into four equal-length segments. For each segment, the sample extractor 210 selects a one minute sample 215 that is centered at the midpoint of the  
30 segment. In this fashion, the sample encoder identifies four, one minute samples 215

that are evenly distributed throughout the source data 105. Together, the samples 215 serve as a proxy for the source data 105. Because the samples 215 include less information than the source data 105, the time required to evaluate the complexity of the samples 215 is less than the time required to evaluate the complexity of the 5 source data 105.

**[0033]** In alternate embodiments, the sample extractor 210 may partition the source data 105 into any number of segments and extract samples from the segments in any manner that generates a representative proxy for the source data 105. Further, the sample extractor 210 may be configured to generate any number 10 and length of the samples 215 in any manner (e.g., meeting a constraint for maximum processing time). In yet other embodiments, the sample extractor 210 is omitted and the source data 105 is processed as a single, full-length sample 215.

**[0034]** The complexity analyzer 220 receives the four samples 215 and configures the encoder 140 to perform fixed QP encodes of the four samples 215 at a resolution 15 of 1080p across four different QPs 225. In operation, the complexity analyzer 220 configures the encoder 140 to execute sixteen encoding tasks 233, thereby generating sixteen constant QP encodes 235. The values of the QPs 225 are experimentally and/or heuristically determined and vetted to provide good coverage across a range of QPs 225 that are anticipated to effectively reduce resource usage 20 without unacceptable degrading quality for any number of complexities. In alternate embodiments, the complexity analyzer 220 may customize the encoder 140 to perform any number of the encoding tasks 233 across any number of QPs 225 and any number of samples 215, in any combination. Further, the complexity analyzer 220 may perform fixed QP encodes for any number and values of resolutions.

**[0035]** To increase the accuracy of the complexity analysis process for the system 25 100, the complexity analyzer 220 leverages the encoder 140 that is included as part of the encoding process for the source data 105. However, in alternate embodiments, the complexity analyzer 220 may configure any available number and type of encoder to generate the constant QP encodes 235.

**[0036]** For each of the constant QP encodes 235, the complexity analyzer 220 then determines the peak signal-to-noise ratio (PSNR)—an objective quality metric based on mean-squared-error—value and calculates the bitrate. In some embodiments, the complexity analyzer 220 configures the encoder 140 to measure the PSNR values of the constant QP encodes 235. In alternate embodiments the complexity analyzer 220 uses a PSNR measurement tool that may or may not be part of the complexity analyzer 220 to measure the PSNR value of each of the constant QP encodes 235.

**[0037]** The complexity analyzer 220 determines the bitrate of each of the constant QP encodes 235 in any technically feasible fashion. In some embodiments, the complexity analyzer 220 performs read operations on log files (generated by the encoder 140 during the encoding process) to identify the bitrate. In other embodiments, the complexity analyzer 220 calculates the bitrate by dividing the size of the constant QP encode 235 (i.e., the file size) by the duration of the constant QP encode 235 (i.e., number of frames divided by the frames per second).

**[0038]** After obtaining the PSNR values and the corresponding bitrates, the complexity analyzer 220 correlates the PSNR values and the corresponding bitrates. In some embodiments, prior to performing correlation operations, the complexity analyzer 220 averages the PSNR values and/or the corresponding bitrates for each of the QPs 255. More specifically, for QP 255<sub>0</sub>, the complexity analyzer 220 averages across the four PSNR values and across the four bitrates obtained for the four samples 115 during the encoding of the constant QP encodes 235 at the QP 255<sub>0</sub>.

**[0039]** As shown, the complexity analyzer 220 generates a PSNR graph 255 that reflects the determined correlation. In operation, the complexity analyzer 220 plots the PSNR values versus the bitrates for the constant QP encodes 235 and then performs curve fitting operations to generate a curve equation. In alternate embodiments, instead of the PSNR graph 255, the complexity analyzer 220 generates an equation that includes a PSNR variable and a bitrate variable based on statistics associated with the constant QP encodes 235. In general, embodiments of the present invention may employ any technically feasible technique and/or any

quality measurement instead of PSNR to determine an estimated relationship between the quality and bitrates of the constant QP encodes 255.

### Assigning Source-Tuned Bitrate Ladder

**[0040]** The bucketing unit 260 evaluates the PSNR graph 255 in conjunction with the complexity buckets 132 and the bitrate ladders 134 to generate the source-tuned bitrate ladder 135. More specifically, the bucketing unit 260 determines a complexity bucketing bitrate based on the PSNR graph 255. The complexity bucketing bitrate is the bitrate at which the PSNR value of the PSNR graph 255 equals a predetermined low distortion threshold. Consequently, the complexity bucketing bitrate is the

estimated bitrate at which the distortion of the encoded data 145 relative to the source data 105 matches the predetermined low distortion threshold. The predetermined low distortion threshold may be assigned using any of a variety of heuristics and experimental techniques that are consistent with the complexity buckets 132 and the bitrate ladders 134.

**[0041]** In general, the bucketing unit 260 may determine the complexity bucketing bitrate in any technically feasible fashion. For example, the bucketing unit 260 may identify the intersection between the curve of the PSNR graph 255 and the line corresponding to a constant PSNR value of the predetermined low distortion threshold. In alternate embodiments the bucketing unit 260 may set the PSNR variable of a PSNR/bitrate equation to the predetermined low distortion threshold and then solve the PSNR/bitrate equation for the value of the bitrate variable.

**[0042]** Subsequently, the bucketing unit 260 compares the complexity bucketing bitrate to the bitrate ranges of the complexity buckets 132 and selects the encompassing complexity bucket 132. The bucketing unit 260 then sets the source-tuned bitrate ladder 135 to the bitrate ladder 134 that corresponds to the selected complexity bucket 132. In this fashion, the source-tuned bitrate ladder 135 varies based on the estimated complexity of the source 105. Advantageously, unlike conventional techniques that rely on a single bitrate ladder, this categorization process optimizes resource/quality tradeoffs across different complexities of the source 105.

[0043] Figure 3 is a conceptual illustration of the complexity buckets 132 and the bitrate ladders 134 implemented by the encoding bitrate ladder selector 130 of Figure 2, according to one embodiment of the present invention. As shown, the complexity buckets 132 include three different classifications based on the bitrate at a low

5 correlation threshold of PSNR value equal to 41 decibels (dB). As persons skilled in the art will recognize, if the PSNR value is greater than 40 dB, then the encoded data 145 is generally considered very low distortion compared to the source data 105. Alternate embodiments may include any number of classifications and any low correlation threshold.

10 [0044] Based on the PSNR graph 255, the bucketing unit 260 determines the complexity bucketing bitrate (R)—the bitrate value of the curve corresponding to a PSNR value of 41 dB. As shown, the complexity buckets 132 include the low complexity bucket 132<sub>1</sub>, the medium complexity bucket 132<sub>2</sub>, and the high complexity bucket 132<sub>3</sub>. If the complexity bucketing bitrate is less than 1750 kilobits per second  
15 (kbps) then the bucketing unit 260 assigns the source data 135 to the low complexity bucket 132<sub>1</sub>. If the complexity bucketing bitrate lies between 1750 kbps and 4300 kbps, then the bucketing unit 260 assigns the source data 135 to the medium complexity bucket 132<sub>2</sub>. If the complexity bucketing bitrate is at least 4300 kbps, then the bucketing unit 260 assigns the source data 135 to the high complexity bucket  
20 132<sub>3</sub>.

[0045] As also shown, each of the complexity buckets 132 is associated with a different one of the bitrate ladders 134. The bitrate / resolution pairs define the rungs of each of the bitrate ladders 134. In addition, each rung includes a profile that specifies the complexity of the algorithm that the encoder 140 employs during the  
25 encoding process. The values of each rung and the number of rungs are tailored for each of the complexity buckets 132. Together, the complexity buckets 132 and the bitrate ladders 134 are designed to optimize tradeoffs between resources and quality improvements. For example, for low complexity sources, maximum perceptible visual quality is achieved at a bitrate of around 3000 kbps. Because encoding at higher  
30 bitrates wastes resources without noticeably increasing visual quality, the highest

rung of the low complexity bitrate ladder  $134_1$  is at 3000 kbps and, consequently, the encoder 140 generates the encoded data 145 with a maximum bitrate of 3000 kbps.

**[0046]** Figure 4 is a conceptual illustration of a peak signal-to-noise ratio (PSNR) curve for a simple source 400 constructed by the complexity analyzer 220 of Figure 2, according to one embodiment of the present invention. As shown, the complexity bucketing bitrate 410 (i.e., the bitrate of the curve at the low correlation threshold of PSNR value equal to 41 dB) is 600 kbps. More specifically, the horizontal dotted line depicts the constant line with PSNR value equal to 41 dB. The horizontal dotted line intersects the PSNR curve for a simple source 400 at the complexity bucketing bitrate 410, shown as an "x." Tracing vertically downwards from the complexity bucketing bitrate 410 (following the vertical dotted line), shows that the complexity bucketing bitrate 410 intersects the bitrate axis at 600 kbps.

**[0047]** Referring back to Figure 3, based on the bitrate of 600 kbps, the bucketing unit 260 assigns the source data 105 to the low complexity bucket  $132_1$  and sets the source-tuned bitrate ladder 135 to the corresponding low complexity bitrate ladder  $134_1$ . Notably, implementing the bitrate ladder  $134_1$  ensures that the highest encoding resolution and bitrate are, respectively, 1080p and 3000 kbps (the highest rung of the low complexity bitrate ladder  $134_1$ ). Advantageously, since higher encoding resolutions and bitrates, such as the 1080p resolution and 4300 kbps bitrate rung of the medium complexity bitrate ladder  $134_1$ , provide no additional quality gain for simple source data 105, limiting the encoding resolution and bitrate conserves resources without impacting quality.

**[0048]** Figure 5 is a conceptual illustration of a peak signal-to-noise ratio (PSNR) curve for a complex source constructed by the complexity analyzer of Figure 2, according to one embodiment of the present invention. As shown, the complexity bucketing bitrate 410 (i.e., the bitrate of the curve at the low correlation threshold of PSNR value equal to 41 dB) is 6000 kbps. More specifically, the horizontal dotted line depicts the constant line with PSNR value equal to 41 dB. The horizontal dotted line intersects the PSNR curve for a complex source 500 at the complexity bucketing bitrate 410, shown as an "x." Tracing vertically downwards from the complexity

bucketing bitrate 410 (following the vertical dotted line), shows that the complexity bucketing bitrate 410 intersects the bitrate axis at 6000 kbps.

**[0049]** Referring back to Figure 3, based on the bitrate of 6000 kbps, the bucketing unit 260 assigns the source data 105 to the high complexity bucket 132<sub>3</sub> and sets the 5 source-tuned bitrate ladder 135 to the corresponding high complexity bitrate ladder 134<sub>3</sub>. Notably, implementing the bitrate ladder 134<sub>3</sub> configures the encoder 140 to encode the source data 105 at a maximum resolution of 1080p and a maximum bitrate of 7500 mbps. Because the source data 105 is relatively complex, such a tradeoff enables noticeable quality improvement using available resources. By 10 contrast, if the source-tuned bitrate ladder 135 were to be the medium complexity bitrate ladder 134<sub>2</sub>, then the encoding bitrate would be unnecessarily limited to 5800 mbps.

**[0050]** Figure 6 is a flow diagram of method steps for selecting and implementing a bitrate ladder while encoding video source data, according to one embodiment of the 15 present invention. Although the method steps are described with reference to the systems of Figures 1-5, persons skilled in the art will understand that any system configured to implement the method steps, in any order, falls within the scope of the present invention. For discussion purposes only, it is assumed in this description of Figure 6 that a low distortion threshold, QPs 225, the bitrate ladders 134, and the 20 complexity buckets 132 are predetermined in any technically feasible fashion.

**[0051]** As shown, a method 600 begins at step 604, where the encoding bitrate ladder selector 130 receives the source data 105, and the sample extractor 210 partitions the source data 105 into N segments, where N is any positive integer. At step 606, for each of the N segments, the sample extractor 210 selects a fixed-length 25 sample 215 centered at the midpoint of the segment. Advantageously, the resulting N samples 215 serve as a proxy for the source data 105, reducing analysis time by limiting the total amount of data that is evaluated to classify the complexity of the source data 105.

**[0052]** The complexity analyzer 220 receives the samples 215 and then 30 configures the encoder 140 to perform fixed QP encodes of the samples 215 at a

fixed resolution across M predetermined, constant QPs 225. At step 608, the complexity analyzer 220 determines a PSNR value for each of the (N \* M) constant QP encodes 235. In some embodiments, the complexity analyzer 220 configures the encoder 140 to measure the PSNR values of the constant QP encodes 235. In 5 general, embodiments of the present invention may replace PSNR with any quality metric, and the subsequently measurements and calculations are modified accordingly.

**[0053]** At step 610, the complexity analyzer 220 determines a corresponding bitrate value for each of the constant QP encodes 235. The complexity analyzer 220 10 determines the bitrate of each of the constant QP encodes 235 in any technically feasible fashion. In some embodiments, the complexity analyzer 220 calculates the bitrate by dividing the size of the constant QP encode 235 (i.e., the file size) by the duration of the constant QP encode 235 (i.e., number of frames divided by the frames per second).

15 **[0054]** At step 614, the complexity analyzer 220 generates the PSNR graph 255 that includes a best-fit curve relating the PSNR values and the corresponding bitrates. In alternate embodiments, instead of the PSNR graph 255, the complexity analyzer 220 generates an equation that includes a PSNR variable and a bitrate variable 20 based on statistics associated with the constant QP encodes 235. In general, embodiments of the present invention may employ any technically feasible technique to determine an estimated relationship between the quality and bitrates of the constant QP encodes 255.

**[0055]** At step 616, the bucketing unit 260 evaluates the PSNR graph 255 and 25 determines the complexity bucketing bitrate 410. Notably, the bucketing unit 260 sets the complexity bucketing bitrate 410 to the value of the bitrate in the PSNR graph 255 where the PSNR value is equal to a predetermined low distortion threshold. Subsequently, the bucketing unit 260 selects the complexity bucket 132 that corresponds to a range of bitrates that includes the complexity bucketing bitrate 410. At step 618, the bucketing unit 260 selects the bitrate ladder 134 that corresponds to 30 the selected complexity bucket 132. At step 620, the bucketing unit 260 sets the

source-tuned bitrate ladder 135 to the selected bitrate ladder 134, and the method 600 terminates.

[0056] Notably, the source-tuned bitrate ladder 135 is tuned to optimize the encoding bitrates and resolution for source data of comparable complexity to the 5 source data 105. For example, referring back to Figure 3, for source data 105 that is relatively simple, the source-tuned bitrate ladder 135 does not include bitrate rungs higher than 3000 kbps that provide no noticeably quality improvement. By contrast, for source data 105 that is relatively complex, the source-tuned bitrate ladder 135 includes a bitrate rung of 7500 kbps that provides increased quality at the expense of 10 resources, such as memory and bandwidth.

[0057] In sum, the disclosed techniques may be used to efficiently select an optimized bitrate ladder (i.e., pairs of bitrates and resolutions) for encoding source data. In operation, a source inspector extracts “N” distributed sample segments from the source data. For each of the sample segments, a complexity analyzer configures 15 an encoder to perform fixed quantization parameter encodes across “M” different values for the quantization parameter. The complexity analyzer then generates an equation for a peak signal-to-noise ratio (PSNR) curve that reflects the relationship between the PSNR and the bitrate for the  $(N * M)$  encoded data segments.

[0058] Subsequently, a bucketing unit sets the PSNR variable of this equation to a 20 predetermined value that represents an acceptable level of distortion, and solves the equation to obtain a bucketing bitrate. Based on the bucketing bitrate the bucketing unit assigns the source data to one of multiple complexity buckets, where each complexity bucket spans a bitrate range that achieves the acceptable level of distortion for source data of a particular complexity. For instance, one bucket may 25 represent simple source data such as cartoons, while another bucket may represent complex source data such as detailed action movies. Notably, each complexity bucket is associated with a bitrate ladder that is empirically determined to optimize encoding source data of the corresponding complexity. .

[0059] Advantageously, sagaciously creating complexity buckets and encoding 30 each source data using the bitrate ladder of the appropriate complexity bucket

optimizes encoding for source data of different complexities. More specifically, because the disclosed techniques tune the bitrate ladder to the complexity of the source, bucketing enables prudent tradeoffs between encoding quality and encoding resources, such as bandwidth and storage. By contrast, conventional encoding processes that rely on a single bitrate ladder often result in encoding that is of an unacceptable low quality or does not effectively exploit opportunities to increase or decrease the use of encoding resources based on judiciously estimated perceptible quality differences.

5 [0060] The descriptions of the various embodiments have been presented for purposes of illustration, but are not intended to be exhaustive or limited to the 10 embodiments disclosed. Many modifications and variations will be apparent to those of ordinary skill in the art without departing from the scope and spirit of the described 15 embodiments.

10 [0061] Aspects of the present embodiments may be embodied as a system, 15 method or computer program product. Accordingly, aspects of the present disclosure may take the form of an entirely hardware embodiment, an entirely software embodiment (including firmware, resident software, micro-code, etc.) or an embodiment combining software and hardware aspects that may all generally be referred to herein as a "circuit," "module" or "system." Furthermore, aspects of the 20 present disclosure may take the form of a computer program product embodied in one or more computer readable medium(s) having computer readable program code embodied thereon.

25 [0001] Any combination of one or more computer readable medium(s) may be utilized. The computer readable medium may be a computer readable signal medium or a computer readable storage medium. A computer readable storage medium may be, for example, but not limited to, an electronic, magnetic, optical, electromagnetic, infrared, or semiconductor system, apparatus, or device, or any suitable combination of the foregoing. More specific examples (a non-exhaustive list) of the computer readable storage medium would include the following: an electrical connection having 30 one or more wires, a portable computer diskette, a hard disk, a random access memory (RAM), a read-only memory (ROM), an erasable programmable read-only

memory (EPROM or Flash memory), an optical fiber, a portable compact disc read-only memory (CD-ROM), an optical storage device, a magnetic storage device, or any suitable combination of the foregoing. In the context of this document, a computer readable storage medium may be any tangible medium that can contain, or store a 5 program for use by or in connection with an instruction execution system, apparatus, or device.

**[0002]** Aspects of the present disclosure are described above with reference to flowchart illustrations and/or block diagrams of methods, apparatus (systems) and computer program products according to embodiments of the disclosure. It will be 10 understood that each block of the flowchart illustrations and/or block diagrams, and combinations of blocks in the flowchart illustrations and/or block diagrams, can be implemented by computer program instructions. These computer program instructions may be provided to a processor of a general purpose computer, special purpose computer, or other programmable data processing apparatus to produce a 15 machine, such that the instructions, which execute via the processor of the computer or other programmable data processing apparatus, enable the implementation of the functions/acts specified in the flowchart and/or block diagram block or blocks. Such processors may be, without limitation, general purpose processors, special-purpose processors, application-specific processors, or field-programmable 50

**[0003]** The flowchart and block diagrams in the Figures illustrate the architecture, functionality, and operation of possible implementations of systems, methods and computer program products according to various embodiments of the present disclosure. In this regard, each block in the flowchart or block diagrams may 20 represent a module, segment, or portion of code, which comprises one or more executable instructions for implementing the specified logical function(s). It should 25 also be noted that, in some alternative implementations, the functions noted in the block may occur out of the order noted in the figures. For example, two blocks shown in succession may, in fact, be executed substantially concurrently, or the blocks may sometimes be executed in the reverse order, depending upon the functionality 30 involved. It will also be noted that each block of the block diagrams and/or flowchart illustration, and combinations of blocks in the block diagrams and/or flowchart

2016243988 17 Dec 2018

illustration, can be implemented by special purpose hardware-based systems that perform the specified functions or acts, or combinations of special purpose hardware and computer instructions.

**[0004]** While the preceding is directed to embodiments of the present disclosure,

5 other and further embodiments of the disclosure may be devised without departing from the basic scope thereof, and the scope thereof is determined by the claims that follow.

**[0005]** As used herein, except where the context requires otherwise the term

'comprise' and variations of the term, such as 'comprising', 'comprises' and

10 'comprised', are not intended to exclude other additives, components, integers or steps.

2016243988 17 Dec 2018

**WHAT IS CLAIMED IS:**

1. A computer-implemented method for selecting a bitrate ladder for encoding source data, the method comprising:
  - 5 selecting a set of parameter values for a quantization parameter;
  - 10 for each parameter value, performing an encoding process that configures an encoder to encode source data to generate a set of encoded data while maintaining the quantization parameter at the parameter value;
  - 15 for each set of encoded data, determining a value of a video quality metric and a corresponding bitrate;
  - 20 deriving a relationship between the video quality metrics and the corresponding bitrates;
  - determining based on the relationship, a bucketing bitrate corresponding to a first video quality metric, wherein the first video quality metric is equal to a predetermined threshold;
  - 15 determining that the bucketing bitrate falls within a bitrate range associated with a first bitrate ladder included in a plurality of bitrate ladders, wherein each bitrate ladder included in the plurality of bitrate ladders is associated with a different bitrate range and includes a different plurality of bitrate-resolution pairs; and
  - 20 selecting the first bitrate ladder for encoding the source data.
2. The computer-implemented method of claim 1, wherein configuring the encoder to encode source data while maintaining the quantization parameter at a given parameter value comprises setting an amount of signal detail to include in the encoded data.
- 25
3. The computer-implemented method of claim 1 or claim 2, wherein selecting the set of parameter values comprises selecting an amount of signal detail that satisfies both an accuracy constraint and a complexity constraint.
- 30
4. The computer-implemented method of any one of the preceding claims, wherein determining the value of a video quality metric comprises identifying a peak

17 Dec 2018

2016243988

signal-to-noise ratio.

5. The computer-implemented method of any one of claim 1 to claim 3, wherein determining the value of a video quality metric comprises calculating a picture quality rating.

6. The computer-implemented method of any one of the preceding claims, wherein determining a corresponding bitrate for a set of encoded data comprises performing a division operation based on a size associated with the set of encoded data and an amount of time associated with the play-out duration of the set of encoded data.

7. The computer-implemented method of any one of the preceding claims, wherein the relationship is a curve, and deriving the relationship comprises applying one or more curve fitting operations to the set of encoded data.

8. The computer-implemented method of any one of the preceding claims, further comprising, prior to configuring the encoder, extracting a plurality of samples from the source data, wherein the encoder generates the sets of encoded data based on signal detail associated with the plurality of samples.

9. A non-transitory computer-readable storage medium including instructions that, when executed by a processing unit, cause the processing unit to select a bitrate ladder for encoding source data by performing the steps of:

25 selecting a set of parameter values for a quantization parameter; for each parameter value, performing an encoding process that configures an encoder to encode source data to generate a set of encoded data while maintaining the quantization parameter at the parameter value;

for each set of encoded data, determining a value of a video quality metric and a corresponding bitrate;

30 deriving a relationship between the video quality metrics and the corresponding bitrates;

determining, based on the relationship, a bucketing bitrate corresponding to a

first video quality metric, wherein the first video quality metric is equal to a predetermined threshold;

5 determining that the bucketing bitrate falls within a bitrate range associated with a first bitrate ladder included in a plurality of bitrate ladders, wherein each bitrate ladder included in the plurality of bitrate ladders is associated with a different bitrate range and includes a different plurality of bitrate-resolution pairs; and

selecting the first bitrate ladder for encoding the source data.

10 10. The non-transitory computer-readable storage medium of claim 9, wherein configuring the encoder to encode source data while maintaining the quantization parameter at a given parameter value comprises setting an amount of signal detail to include in the encoded data.

15 11. The non-transitory computer-readable storage medium of claim 9 or claim 10, further comprising, prior to configuring the encoder, setting the predetermined threshold to correspond to a maximum acceptable amount of distortion.

20 12. The non-transitory computer-readable storage medium of any one of claim 9 to 11, wherein determining the value of a video quality metric comprises identifying a peak signal-to-noise ratio.

25 13. The non-transitory computer-readable storage medium of any one of claim 9 to 11, wherein determining the value of a video quality metric comprises performing one or more read operations on a log file generated by the encoder.

14. The non-transitory computer-readable storage medium of any one of claim 9 to 13, wherein determining a corresponding bitrate for a set of encoded data comprises performing one or more read operations on a log file generated by the encoder.

30 15. The non-transitory computer-readable storage medium of any one of claim 9 to claim 14, wherein the relationship is an equation.

17 Dec 2018

2016243988

16. The non-transitory computer-readable storage medium of any one of claim 9 to claim 15, further comprising, prior to configuring the encoder, extracting a plurality of clips from the source data, wherein the encoder generates the sets of encoded data based on signal detail associated with the plurality of clips and each clip represents a predetermined length of time.

17. A system configured select a bitrate ladder for encoding source data, the system comprising:

10 a memory storing instructions;

15 a processor executing instructions to:

select a set of parameter values for a quantization parameter;

for each parameter value, perform an encoding process that configures

an encoder to encode source data to generate a set of encoded data while maintaining the quantization parameter at the parameter value;

for each set of encoded data, determine a value of a video quality metric and a corresponding bitrate;

derive a relationship between the video quality metrics and the corresponding bitrates;

20 determine, based on the relationship, a bucketing bitrate range

corresponding to a first video quality metric, wherein the first video quality metric is equal to a predetermined threshold;

determine that the bucketing bitrate falls within a bitrate range

associated with a first bitrate ladder included in a plurality of

25 bitrate ladders, wherein each bitrate ladder included in the

plurality of bitrate ladders is associated with a different bitrate

range and includes a different plurality of bitrate-resolution pairs;

and

30 select the first bitrate ladder for encoding the source data.

18. The system of claim 17, wherein configuring the encoder to encode source data while maintaining the quantization parameter at a given parameter value comprises setting an amount of signal detail to include in the encoded data.

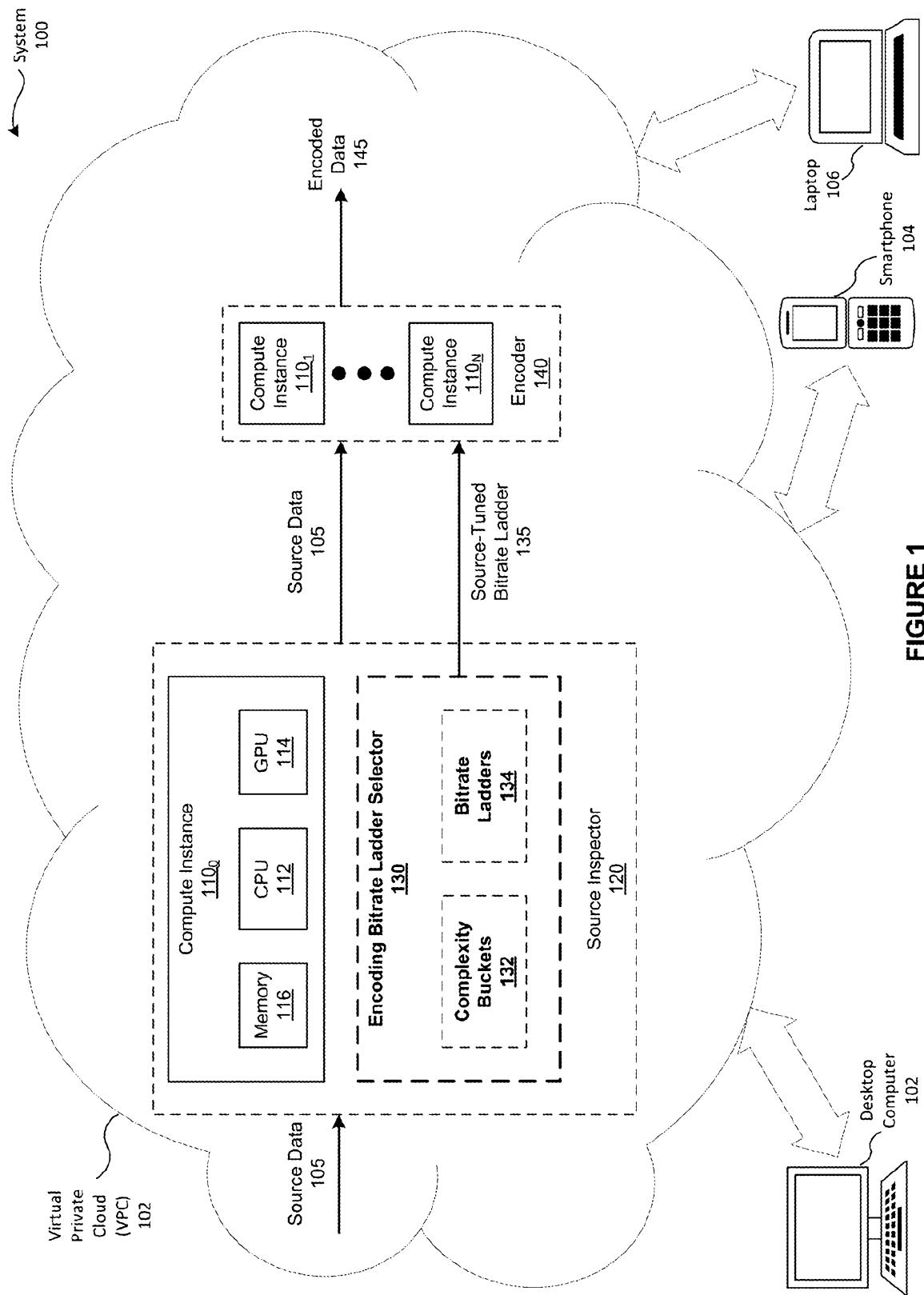


FIGURE 1

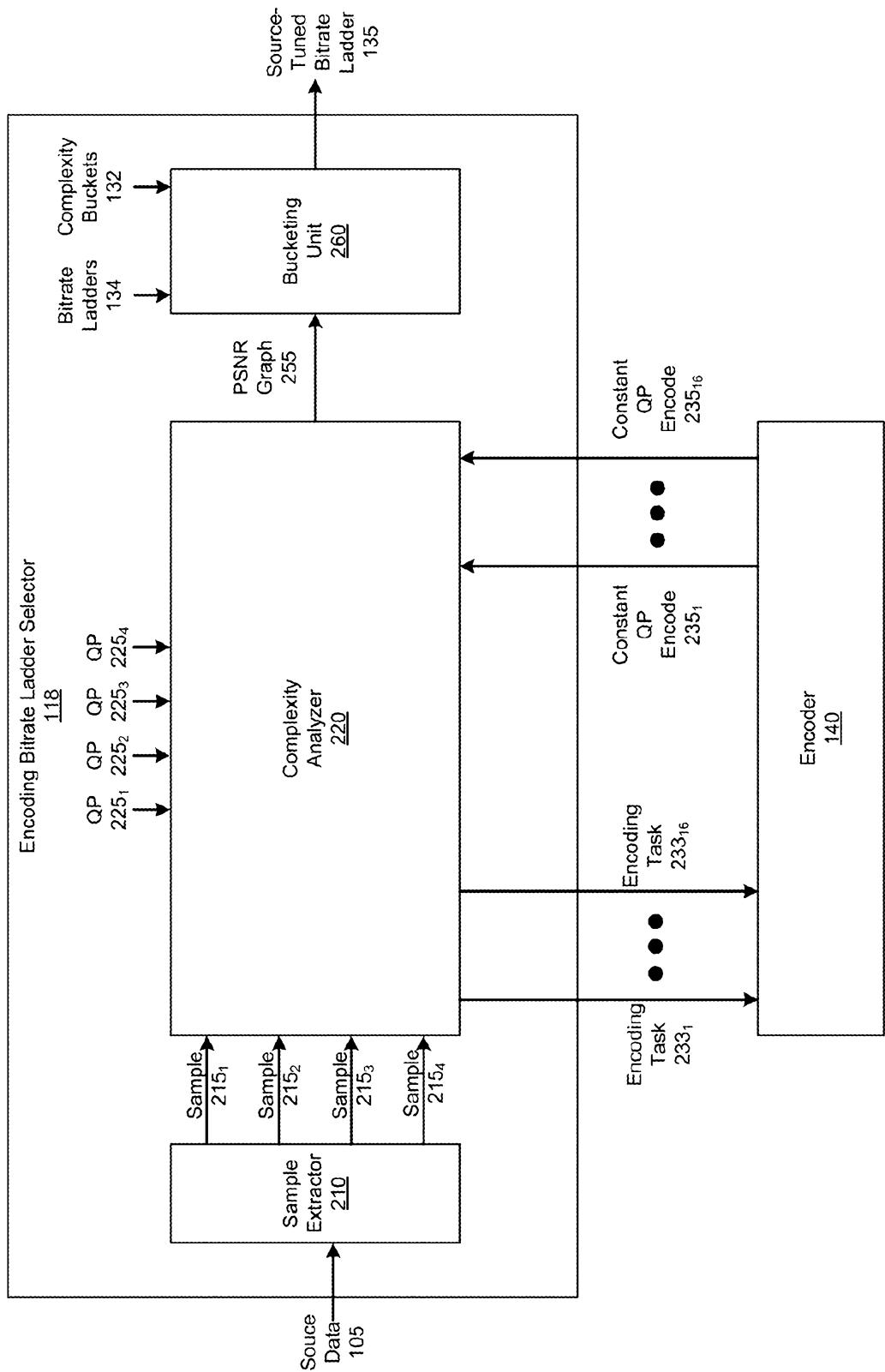


FIGURE 2

**Bitrate at PSNR of 41 Classification**

Bitrate at PSNR of 41	Classification
$R < 1750$	Bucket 1 Low Complexity
$1750 \leq R < 4300$	Bucket 2 Medium Complexity
$4300 \leq R$	Bucket 3 High Complexity

**Complexity Bucket 132<sub>1</sub>**

Bitrate	Resolution	Profile
375	480	Base
560	480	Base
750	480	Base
1050	720	Main
1750	1080	High
2350	1080	High
3000	1080	High
4300		
5800		
7500		

**Complexity Bucket 132<sub>2</sub>**

Bitrate	Resolution	Profile
1080	720	Main

**Complexity Bucket 132<sub>3</sub>**

Bitrate	Resolution	Profile
1080	480	Base
1080	720	Main

**Bucket 1**

Bitrate	Resolution	Profile
375	480	Base
560	480	Base
750	480	Base
1050	720	Main
1750	1080	High
2350	1080	High
3000	1080	High
4300		
5800		
7500		

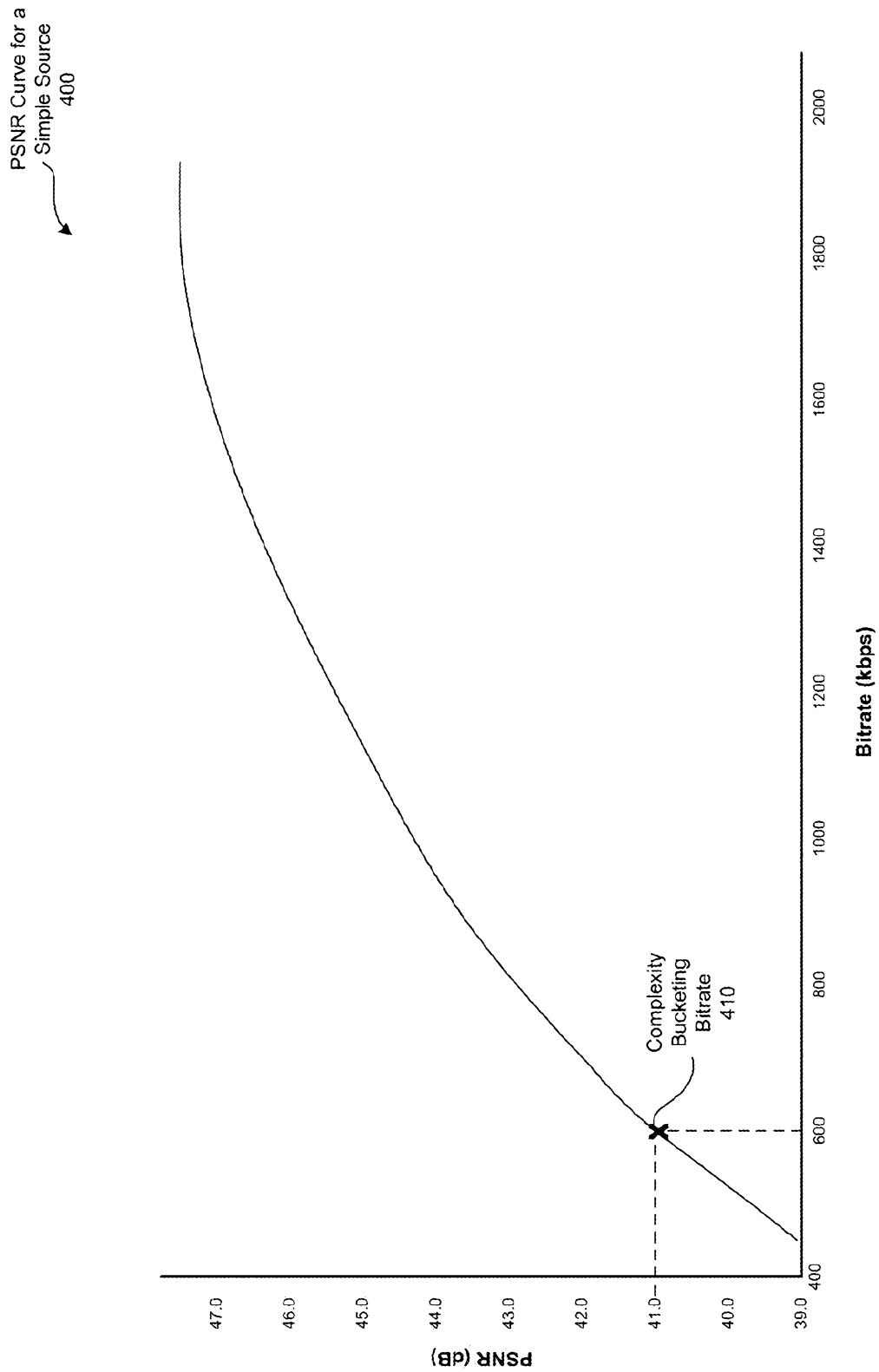
**Bucket 2**

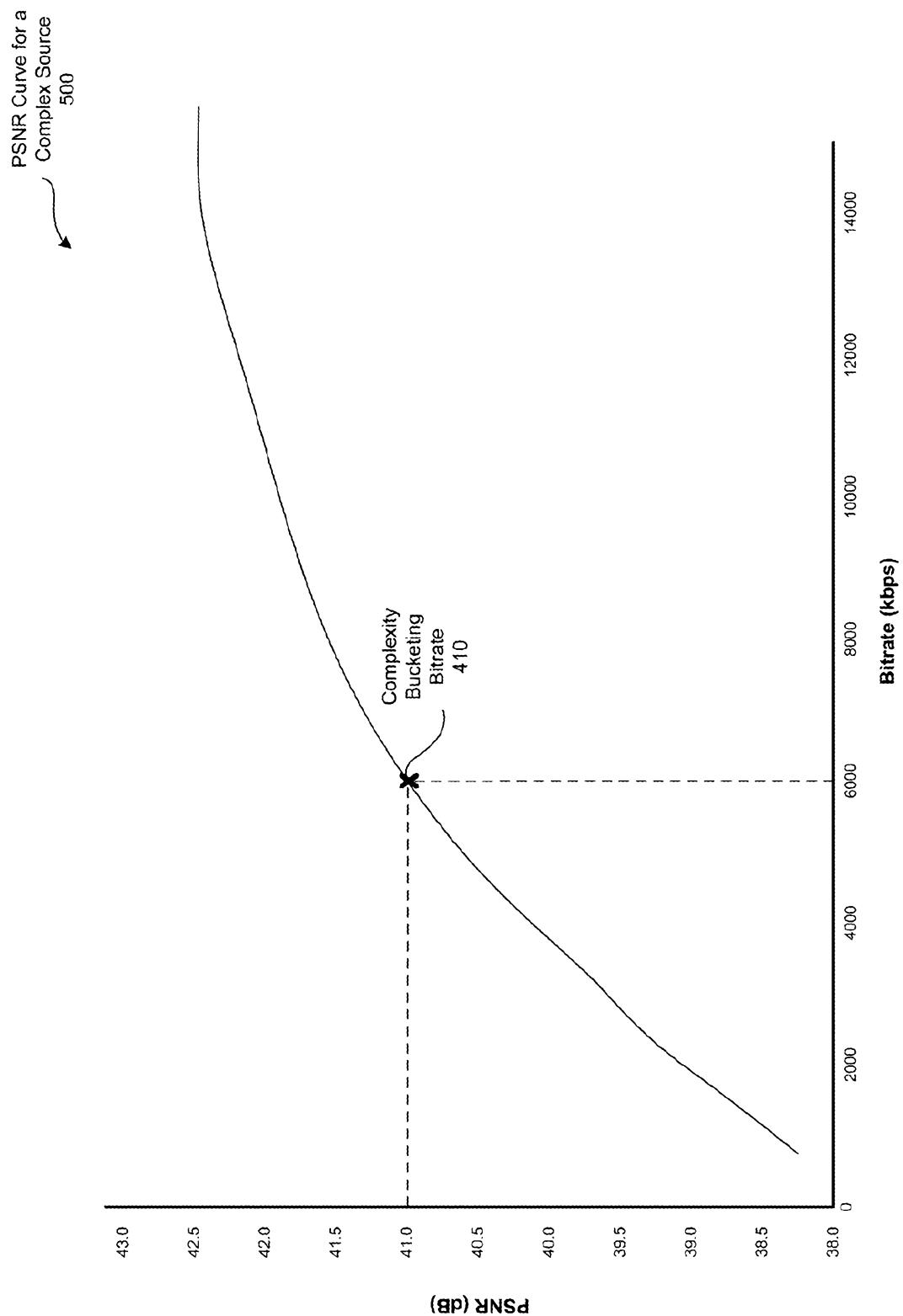
Bitrate	Resolution	Profile
1080	480	Base
1080	720	Main

**Bucket 3**

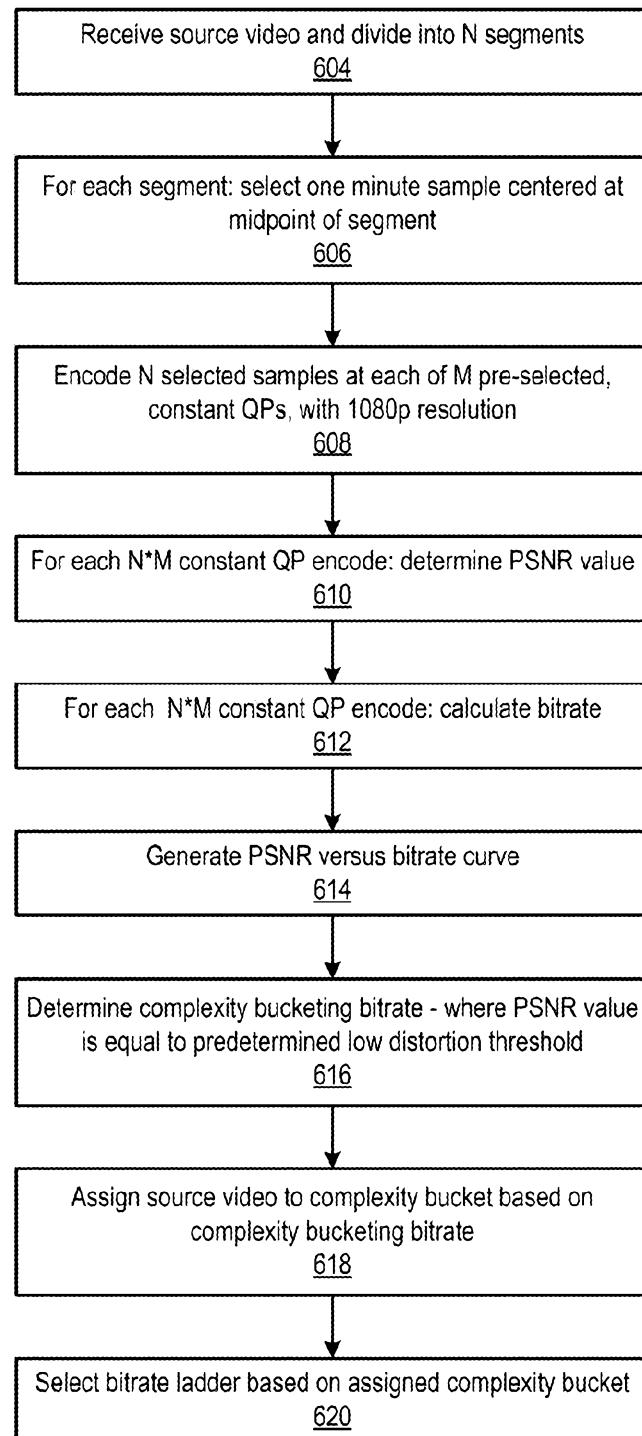
Bitrate	Resolution	Profile
1080	480	Base
1080	720	Main

**FIGURE 3**

**FIGURE 4**

**FIGURE 5**

600

**FIGURE 6**