

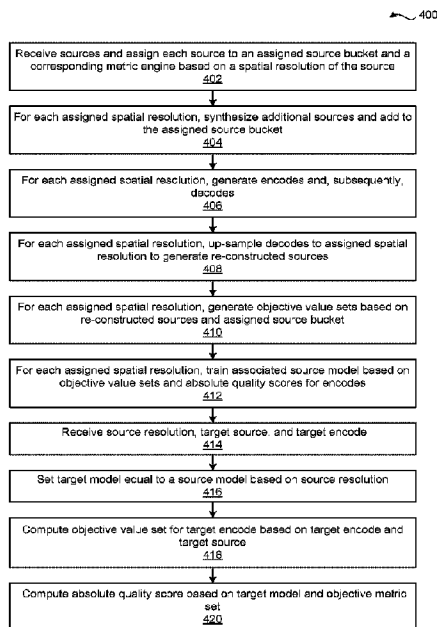


(86) Date de dépôt PCT/PCT Filing Date: 2017/12/06
(87) Date publication PCT/PCT Publication Date: 2018/06/21
(45) Date de délivrance/Issue Date: 2023/09/19
(85) Entrée phase nationale/National Entry: 2019/06/05
(86) N° demande PCT/PCT Application No.: US 2017/064975
(87) N° publication PCT/PCT Publication No.: 2018/111655
(30) Priorités/Priorities: 2016/12/12 (US62/432,870);
2017/10/12 (US15/782,586); 2017/10/12 (US15/782,590)

(51) Cl.Int./Int.Cl. *H04N 19/154* (2014.01),
H04N 19/00 (2014.01), *H04N 19/59* (2014.01)
(72) Inventeurs/Inventors:
LI, ZHI, US;
AARON, ANNE, US;
MOORTHY, ANUSH, US;
BAMPIS, CHRISTOS, US
(73) Propriétaire/Owner:
NETFLIX, INC., US
(74) Agent: DEETH WILLIAMS WALL LLP

(54) Titre : TECHNIQUES A SOURCE CONSTANTE DESTINEES A LA PREDICTION DE QUALITE VIDEO
PERCEPTUELLE ABSOLUE

(54) Title: SOURCE-CONSISTENT TECHNIQUES FOR PREDICTING ABSOLUTE PERCEPTUAL VIDEO QUALITY



(57) **Abrégé/Abstract:**

In various embodiments, a perceptual quality application computes an absolute quality score for encoded video content. In operation, the perceptual quality application selects a model based on the spatial resolution of the video content from which the encoded video content is derived. The model associates a set of objective values for a set of objective quality metrics with an absolute quality score. The perceptual quality application determines a set of target objective values for the objective quality metrics based on the encoded video content. Subsequently, the perceptual quality application computes the absolute quality score for the encoded video content based on the selected model and the set of target objective values. Because the absolute quality score is independent of the quality of the video content, the absolute quality score accurately reflects the perceived quality of a wide range of encoded video content when decoded and viewed.

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

(19) World Intellectual Property
Organization
International Bureau

(43) International Publication Date
21 June 2018 (21.06.2018)



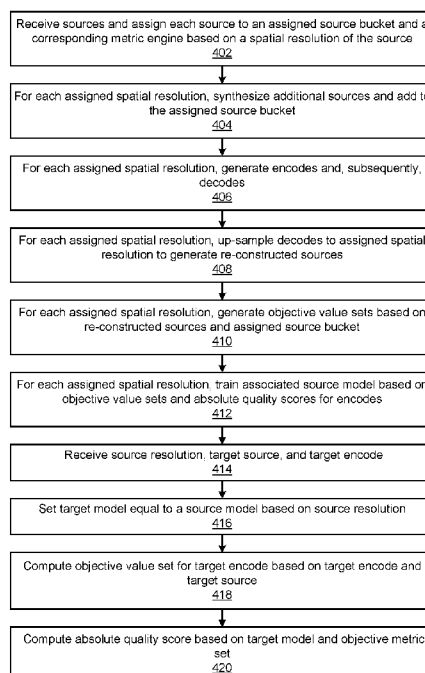
(10) International Publication Number
WO 2018/111655 A1

- (51) **International Patent Classification:**
H04N 19/154 (2014.01) *H04N 19/00* (2014.01)
H04N 19/59 (2014.01)
- (21) **International Application Number:**
PCT/US2017/064975
- (22) **International Filing Date:**
06 December 2017 (06.12.2017)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
62/432,870 12 December 2016 (12.12.2016) US
15/782,586 12 October 2017 (12.10.2017) US
15/782,590 12 October 2017 (12.10.2017) US
- (71) **Applicant:** NETFLIX, INC. [US/US]; 100 Winchester Circle, Los Gatos, California 95032 (US).
- (72) **Inventors:** LI, Zhi; 100 Winchester Circle, Los Gatos, California 95032 (US). AARON, Anne; 100 Winchester Circle, Los Gatos, California 95032 (US). MOORTHY, Anush; 100 Winchester Circle, Los Gatos, California 95032 (US). BAMPIS, Christos; 100 Winchester Circle, Los Gatos, California 95032 (US).
- (74) **Agent:** CAREY, John C. et al.; Arteris 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, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN,

(54) **Title:** SOURCE-CONSISTENT TECHNIQUES FOR PREDICTING ABSOLUTE PERCEPTUAL VIDEO QUALITY

FIGURE 4

400



(57) **Abstract:** In various embodiments, a perceptual quality application computes an absolute quality score for encoded video content. In operation, the perceptual quality application selects a model based on the spatial resolution of the video content from which the encoded video content is derived. The model associates a set of objective values for a set of objective quality metrics with an absolute quality score. The perceptual quality application determines a set of target objective values for the objective quality metrics based on the encoded video content. Subsequently, the perceptual quality application computes the absolute quality score for the encoded video content based on the selected model and the set of target objective values. Because the absolute quality score is independent of the quality of the video content, the absolute quality score accurately reflects the perceived quality of a wide range of encoded video content when decoded and viewed.

[Continued on next page]



WO 2018/111655 A1

WO 2018/111655 A1

HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, 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)*)

SOURCE-CONSISTENT TECHNIQUES FOR PREDICTING ABSOLUTE PERCEPTUAL VIDEO QUALITY

[0001]

5

10

15 **BACKGROUND OF THE INVENTION**

Field of the Invention

[0002] Embodiments of the present invention relate generally to video technology and, more specifically, to source-consistent techniques for predicting absolute perceptual video quality.

20 **Description of the Related Art**

[0003] Efficiently and accurately encoding source video content is critical for real-time delivery of high-quality video content. Because of variations in encoded video content quality, it is desirable to implement quality controls to ensure that the visual quality of decoded video content derived from the encoded video content is acceptable.

25 Manually verifying the visual quality of decoded video content is prohibitively time consuming. Accordingly, automated video content quality assessment is often an integral part of the encoding and streaming infrastructure. For example, automated video quality assessment can be employed when evaluating encoders and fine-tuning streaming bitrates to optimize encoded video content quality.

30 [0004] Some approaches to automatically assessing the quality of encoded video content involve computing scores for a perceptive quality metric based on subjective viewer ratings of the visual quality of decoded video content derived from encoded

video content. For explanatory purposes, a discrete portion of source video content is referred to herein as a “source,” the encoded video content derived from a source is referred to herein as an “encode,” and the decoded video content derived from the encoded video content is referred to herein as a “decode.” One limitation of

5 computing and evaluating scores for perceptive quality metrics is that the metrics typically estimate the quality of the decode relative to the quality of the source. However, because the quality of sources may differ, the computed scores of the perceptive quality metrics do not necessarily reflect human perception of the quality of encodes in a comprehensive fashion.

10 [0005] For example, suppose that the source “A” has a spatial resolution of 480p and appears blurry to viewers, while the source “B” has a spatial resolution of 1080p and appears sharp to viewers. Further, suppose that the computed scores for both the encode “EA” derived from the source “A” and the encode “EB” derived from the source “B” are 100. Finally, suppose that a score of 100 indicates that the visual
15 quality of a decode derived from an encode is perceptively identical to the visual quality of the source from which the encode was derived. In apparent contradiction to the scores being identical, the decode derived from the encode “EA” would appear blurry to viewers, while the decode derived from the encode “EB” would appear sharp to viewers.

20 [0006] Another limitation of computing and evaluating perceptive quality metrics is that the subjective viewer ratings are typically obtained based on viewings of decodes via a particular type of device (e.g., a television). For example, suppose that a source is encoded at 2000 kilobits per second (Kbps) to generate an encode “EA,” and the source is encoded at 3000 Kbps to generate a different encode “EB.” Because the
25 decode derived from the encode EA appears inferior to the decode derived from the encode EB when viewed on the television, the score of the encode EA is lower than the score of the encode EB. However, despite the difference in scores, the decode derived from the encode EA and the decode derived from the encode EB could appear perceptually indistinguishably when viewed on a phone.

30 [0007] As the foregoing illustrates, what is needed in the art are more effective techniques for predicting the perceived quality of decoded video content derived from encoded video content.

SUMMARY OF THE INVENTION

[0008] One embodiment of the present invention sets forth a computer-implemented method for predicting absolute video quality. The method includes selecting a first model based on a first spatial resolution of first video content, where the first model is
5 included in a plurality of models and associates a set of objective values for a set of objective quality metrics with an absolute quality score; computing a first set of values for the set of objective quality metrics based on first encoded video content derived from the first video content; and computing a first absolute quality score for the first encoded video content based on the first model and the first set of values, where the
10 first absolute quality score is aggregated with one or more other absolute quality scores in evaluating the visual quality of streamed media content.

[0009] One advantage of the disclosed techniques for estimating absolute video quality is that the absolute quality score robustly predicts perceived video quality across a diverse range of encoded video content. In particular, the techniques
15 compute the absolute quality score based on a model that inherently reflects a spatial resolution of the source video content from which the encoded video content is derived. Consequently, the perceptual quality score allows meaningful comparison and effective optimization operations involving the encoded video content irrespective of the spatial resolution of the source video content.

20 **BRIEF DESCRIPTION OF THE DRAWINGS**

[0010] So that the manner in which the above recited features of the present invention 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
25 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 effective embodiments.

[0011] Figure 1 is a conceptual illustration of a perceptual quality modeling system configured to implement one or more aspects of the present invention;

30 [0012] Figure 2 is a more detailed illustration of the metric engine of Figure 1, according to various embodiments of the present invention;

[0013] Figure 3 is a more detailed illustration of the perceptual quality calculator of Figure 1, according to various embodiments of the present invention;

[0014] Figure 4 is a flow diagram of method steps for computing an absolute quality score for encoded video content viewed on a base viewing device, according to
5 various embodiments of the present invention; and

[0015] Figure 5 is a flow diagram of method steps for computing an absolute quality score for encoded video content viewed on a target viewing device, according to various embodiments of the present invention.

DETAILED DESCRIPTION

10 [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 more of these specific details.

[0017] In sum, the disclosed techniques may be used to efficiently and reliably predict
15 an absolute quality score for encodes derived from sources. Initially, for each of any number of different spatial resolutions, a training subsystem trains a corresponding source model based on training sources having the spatial resolution. For an encode derived from a source having a particular spatial resolution, the source model corresponding to the spatial resolution associates an objective value set for the
20 encode and absolute quality score for the encode viewed on a base viewing device. For each of any number of additional device types, a conversion subsystem generates a corresponding device equation based on a human-videos viewings of training encodes on the base viewing device and human-video viewings of the training encodes on an additional viewing device having the additional device type.

25 [0018] To determine the absolute quality score for a target encode, a perceptual quality calculator sets a target model equal to a source model associated with the spatial resolution of the source from which the target encode was derived. The perceptual quality calculator then computes a base quality score associated with the target encode based on the target model and a set of objective metrics for the target
30 encode. If the target device type associated with the viewing device on which the target encode is to be viewed is equal to the base device type, then the perceptual

quality calculator sets the absolute quality score associated with the encode equal to the base quality score associated with the encode. If, however, the target device type does not equal the base device type, then the perceptual quality calculator sets a target device equation equal to the device equation associated with the target device type. The perceptual quality calculator then computes the absolute quality score for the target encode based on the target device equation and the base quality score for the target encode. Advantageously, the absolute quality score for a target encode on a target viewing device accurately reflects both the resolution of the source from which the encode was derived and the target viewing device.

10 **Perceptual Quality Modeling System Overview**

[0019] Figure 1 is a conceptual illustration of a perceptual quality modeling system 100 configured to implement one or more aspects of the present invention. As shown, the perceptual quality modeling system 100 includes, without limitation, a cloud (i.e., encapsulated shared resources, software, data, etc.) 102. The cloud 102 may be connected to any number and type of devices capable of transmitting input data and/or displaying video content. Examples of such devices include, without limitation, game consoles, smart phones, smart televisions, laptops, tablets, and desktop computers. Devices that are capable of displaying video content are also referred to herein as “viewing devices.” For explanatory purposes, multiple instances of like objects are denoted with reference numbers identifying the object and parenthetical numbers identifying the instance where needed.

[0020] The cloud 102 includes, without limitation, any number and type of compute instances 110. The cloud 102 receives input user information from an input device (e.g., a laptop), one or more computer instances 110 operate on the user information, and the cloud 102 transmits processed information to the user. The cloud 102 conveys output information to the user via display capabilities of any number of devices, such as a conventional cathode ray tube, liquid crystal display, light-emitting diode, or the like.

[0021] In alternate embodiments, the perceptual quality modeling system 100 may include any distributed computer system instead of the cloud 102. In other embodiments, the perceptual quality modeling system 100 does not include the cloud 102 and, instead, the perceptual quality modeling system 100 includes a single

computing unit that implements any number of processing units (e.g., central processing units and/or graphical processing units in any combination).

- [0022]** As shown, each of the compute instances 110 includes, without limitation, a processor 112 and a memory 116. The processor 112 may be any instruction execution system, apparatus, or device capable of executing instructions. For example, the processor 112 could comprise a central processing unit (CPU), a graphics processing unit (GPU), a controller, a microcontroller, a state machine, or any combination thereof. The memory 116 stores content, such as software applications and data, for use by the processor 112 of the compute instance 110.
- [0023]** The memory 116 may be one or more of a readily available memory, such as random access memory (RAM), read only memory (ROM), floppy disk, hard disk, or any other form of digital storage, local or remote. In some embodiments, a storage (not shown) may supplement or replace the memory 116. The storage may include any number and type of external memories that are accessible to the processor 112. For example, and without limitation, the storage may include a Secure Digital Card, an external Flash memory, a portable compact disc read-only memory (CD-ROM), an optical storage device, a magnetic storage device, or any suitable combination of the foregoing.
- [0024]** In general, the compute instances 110 included in the cloud 102 are configured to implement one or more applications and/or one or more subsystems of applications. For explanatory purposes only, each application and each subsystem is depicted as residing in the memory 116 of a single compute instance 110 and executing on a processor 112 of the single compute instance 110. However, as persons skilled in the art will recognize, the functionality of each application and subsystem may be distributed across any number of other applications and/or subsystems that reside in the memories 116 of any number of compute instances 110 and execute on the processors 112 of any number of compute instances 110 in any combination. Further, the functionality of any number of applications and/or subsystems may be consolidated into a single application or subsystem.
- [0025]** In general, the perceptual quality modeling subsystem 100 provides automated video content quality assessment for encoded video content. The video content quality assessment may be used to implement quality controls associated with an

encoding and streaming infrastructure. For example, to ensure an acceptable viewing experience, the quality of encoded video content and, indirectly, the caliber of the elements included in the encoding and streaming infrastructure are typically evaluated at various points in the design and delivery process using quality metrics. The values
5 for the quality metrics are then used to guide the development of applications (e.g., encoders) and the real-time optimization of content delivery, such as stream-switching algorithms that are quality-aware.

[0026] Some conventional approaches to automatically assessing the quality of encoded video content, involve computing scores for a conventional perceptive
10 quality metric based on subjective viewer ratings of the visual quality of decoded video content derived from encoded video content. For explanatory purposes, a discrete portion of source video content is referred to herein as a “source,” the encoded video content derived from a source is referred to herein as an “encode,” and the decoded video content derived from the encoded video content is referred to
15 herein as a “decode.” One limitation of computing and evaluating scores for conventional perceptive quality metrics is that the metrics typically estimate the quality of the decode relative to the quality of the source. However, because the quality of sources may differ, the computed scores of the conventional perceptive quality metrics do not necessarily reflect human perception of the quality of encodes in a
20 comprehensive fashion.

[0027] For example, suppose that the source “A” has a spatial resolution of 480p and appears blurry to viewers, while the source “B” has a spatial resolution of 1080p and appears sharp to viewers. Further, suppose that the scores for both the encode “EA” derived from the source “A” and the encode “EB” derived from the source “B” are 100.
25 Finally, suppose that a conventional score of 100 indicates that the visual quality of an decode derived from an encode is perceptively identical to the visual quality of the source from which the encode was derived. In apparent contradiction to the scores being identical, the decode derived from the encode “EA” would appear blurry to viewers, while the decode derived from the encode “EB” would appear sharp to
30 viewers.

[0028] Another limitation of computing and evaluating conventional perceptive quality metrics is that the subjective viewer ratings are typically obtained based on viewings

of decodes via a particular type of viewing device (e.g., a smart television). For example, suppose that a source is encoded at 2000 kilobits per second (Kbps) to generate an encode “EA,” and the source is encoded at 3000 Kbps to generate a different encode “EB.” Because the decode derived from the encode EA appears

5 inferior to the decode derived from the encode EB when viewed on the smart television, the score of the encode EA is significantly lower than the score of the encode EB. However, despite the difference in scores, the decode derived from encode EA and the decode derived from encode EB could appear perceptually indistinguishably when viewed on a smart phone with a relatively small, low resolution

10 display.

Instituting an Absolute Perceptual Quality Metric

[0029] For these reasons, one or more of the compute instances 110 in the perceptual quality model system 100 implement techniques to institute an absolute perceptual quality metric. As shown, the perceptual quality model system 100 includes, without

15 limitation, a subjective metric subsystem 140, any number of training subsystems 150, any number of conversion subsystems 170, and a perpetual quality calculator 160. Each of the subjective metric subsystem 140, the training subsystems 150, the conversion subsystems 170, and the perpetual quality calculator 160 reside in the memory 116 associated with one or more of the compute instances 110 and execute

20 on the processor 112 associated with one or more of the compute instances 110.

[0030] In general, the perceptual quality model system 100 operates in two different phases. A training phase is depicted with non-bold lines and boxes, and a prediction phase is depicted with bold lines and boxes. In the training phase, each of the training subsystems 150 generates a source model 155 associated with a different

25 source resolution 180, and each of the conversion subsystems 170 generates a device equation 175 associated with a different device type. The source resolution 180 is a spatial resolution of a source from which an encode is derived. The device type is a type associated with any number of viewing devices having similar viewing-related properties. For explanatory purposes only, viewing a decoded version of an encode is also referred to herein as “viewing the encode.”

30

[0031] Each of the source models 155 associates an objective value set 165 for an encode and an absolute quality score 195 for the encode viewed on a base viewing

device (not shown in Figure 1). The absolute quality score 195 of a particular encode viewed on a base viewing device is also referred to herein as “the base quality score” of the encode. The objective value set 165 is a set of values for a set of objective metrics. In a complementary fashion, each of the device equations 175 specifies a conversion between the base quality score of an encode and the absolute quality score 195 of the encode viewed on a target viewing device having the associated device type.

[0032] Subsequently, in the prediction phase, a perceptual quality calculator 160 computes the absolute quality score 195 for a target encode 185 based on the target encode 185, a target source 187, the source resolution 180, and a target device type 190. The target source 187 is the source from which the target encode 185 is derived. The target device type 190 specifies a device type that characterizes the target viewing device on which the target encode 185 is to be viewed. The perceptual quality calculator 160 is also referred to herein as the perceptual quality application.

[0033] In operation, the perceptual quality calculator 160 selects one of the source models 155 based on the source resolution 180 as the target model. The perceptual quality calculator 160 then computes the objective value set 165 associated with the target encode 185 based on the target encode 185 and the target source 187. The perceptual quality calculator 160 may compute the objective value set 165 in any technically feasible fashion. Subsequently, the perceptual quality calculator 160 computes the base quality score for the target encode 185 based on the target model and the objective value set 165 associated with the target encode 185.

[0034] If the target device type 190 equals the base device type, then the perceptual quality calculator 160 sets the absolute quality score 195 equal to the base quality score. If, however, the target device type 190 is not equal to the base device type, then the perceptual quality calculator 160 selects the device equation 175 associated with the target device type 190 as the target device equation. The perceptual quality calculator 160 computes the absolute quality score 195 based on the target device equation and the base quality score.

[0035] For explanatory purposes only, the perceptual quality modeling system 100 is configured to differentiate between three different source resolutions 180 and two different target device types 190 (one base device type and one additional device

type). In alternate embodiments, the perceptual quality modeling system 100 may differentiate between any number of source resolutions 180 (including zero) and any number of target device types 190 (including zero) in any combination. For instance, in some alternate embodiments, the perceptual quality model system 100

- 5 differentiates between four different source resolutions 180 (e.g., 480p, 720p, 1080p, and 4K). In the same or other alternate embodiments, the target device type 190 may specify any number of characteristics associated with any number of viewing devices. For example, the target device type 190 could specify a spatial resolution associated with the viewing device instead of or in addition to a type of the target viewing device.

- 10 **[0036]** As shown, the subjective metric subsystem 140 includes, without limitation, three metric engine 142 (i.e., three instances of the metric engine 142). Each of the metric engines 142 is associated with an assigned spatial resolution (not shown in Figure 1) and a different assigned source bucket (not shown in Figure 1). The metric engine 142(1) is associated with the assigned spatial resolution of 480p, the metric
15 engine 142(2) is associated with the assigned spatial resolution of 720p, and the metric engine 143(3) is associated with the assigned spatial resolution of 1080p.

- [0037]** During the training phase, upon receiving any number of sources, the subjective metric subsystem 140 assigns each of the sources to one of the assigned source buckets based on the spatial resolution of the source. The subjective metric
20 subsystem 140 assigns sources having a spatial resolution equal to 480p to the assigned source bucket associated with the metric engine 142(1). The subjective metric subsystem 140 assigns sources having a spatial resolution equal to 720p to the assigned source bucket associated with the metric engine 142(2). The subjective metric subsystem 140 assigns sources having a spatial resolution equal to 1080p to
25 the assigned source bucket associated with the metric engine 142(3).

- [0038]** As described in greater detail in conjunction with Figure 2, each metric engine 142(i) derives encodes from the sources included in the assigned source bucket, generates an objective source bucket 135(i), and generates a subjective source bucket 125(i). The objective source bucket 135(i) includes an objective value set 165
30 for each encode generated by the metric engine 142(i). The objective value set 165 is a set of values for the set of objective metrics. Examples of objective metrics include a visual information fidelity (VIF), a detail loss measure (DLM) an additive

impairment measure (ADM), a visual information fidelity (VIF), a blind/referenceless image spatial quality evaluator (BRISQUE), and a temporal information (TI). The metric engine 142(i) may compute the objective value sets 165 included in the objective source bucket 135(i) in any technically feasible fashion.

- 5 **[0039]** The subjective source bucket 125(i) includes an absolute quality score 195 for each encode generated by the metric engine 142(i). The absolute quality score 195 is derived from values for the absolute perceptual quality metric assigned by humans after viewing the encodes generated by the metric engine 142(i) on a base viewing device (e.g., a smart television). The human-observed values may be obtained via
 10 any type of absolute rating process. For instance, the human-observed values could be determined via an absolute category rating testing process. For each encode, the absolute quality score 195 may be computed in any technically feasible fashion and by any software application. For example, the metric engine 142(i) could compute the mean opinion score (MOS) based on the human-observed values for the absolute
 15 perceptual quality metric, and then set the absolute quality score 195 equal to the MOS.

- [0040]** The subjective metric subsystem 140 also generates a base subjective bucket 142 that includes the absolute quality scores 195 derived from values for the absolute perceptual quality metric assigned by humans after viewing the unique encodes on
 20 the base viewing device. For explanatory purposes only, “unique encodes” refer to the set of encodes generated by the metric engines 142 after any duplicate encodes are removed. Further, for each additional device type, the subjective metric subsystem 140 generates an additional subjective bucket 144 that includes the absolute quality scores 195 derived from values for the absolute perceptual quality
 25 metric assigned by humans after viewing the unique encodes on an additional viewing device associated with the additional device type. As shown, the subjective metric subsystem 140 generates one additional subjective bucket 144(1). In alternate embodiments, the subjective metric subsystem 140 may generate any number of additional subjective buckets 144.

- 30 **[0041]** Each of the training subsystems 150(i) receives the objective source bucket 135(i) and the subjective source bucket 125(i) generated by an associated metric engine 142(i), and generates the source model 155(i). Accordingly, the training

subsystem 150(i) and the source model 155(i) are associated with the assigned spatial resolution that characterizes the sources assigned to the metric engine 142(i). As shown, the training subsystem 150(1) and the source model 155(1) are associated with the spatial resolution of 480p. The training subsystem 150(2) and the source model 155(2) are associated with the spatial resolution of 720p. The training subsystem 150(3) and the source model 155(3) are associated with the spatial resolution of 1080p.

[0042] In operation, the training subsystem 150 generates the source model 155 via a machine learning algorithm trained on the objective source bucket 135 and the subjective source bucket 125. More specifically, each of the objective value sets 165 included in the the objective source bucket 135 is associated with a particular encode and, consequently, a corresponding absolute quality score 195 included in the subjective source bucket 125. For each of the objective value sets 165, the training subsystem 150 trains the source model 155 to associate the objective value set 165 and the corresponding absolute quality score 195. The training subsystem 150 may implement any type of machine learning algorithm in any technically feasible fashion. For example, the training subsystem 150 could implement a support vector machine algorithm, an artificial neural network algorithm, or a random forest algorithm, to name a few.

[0043] For each additional device type, an instance of the conversion subsystem 170(i) receives the base subjective bucket 172 and the additional subjective bucket 174(i), and generates the device equation 175(i). The device equation 175(i) defines a relationship between the base quality score of an encode and the absolute quality score 195 for the encode viewed on a viewing device characterized by the associated additional device type. The conversion subsystem 170 may determine any type of relationship in any technically feasible fashion.

[0044] For instance, in some embodiments, the conversion subsystem 170 generates a scatter plot, where each point included in the scatter plot is associated with a different encode. The horizontal coordinate of a particular point represents the absolute quality score 195 for the associated encode that is included in the base subjective bucket 172. The vertical coordinate of the point represents the absolute quality score 195 for the associated encode that is included in the associated

additional subjective bucket 174. The conversion subsystem 170 then performs curve fitting operations on the scatter plot to generate a polynomial. The polynomial expresses the absolute quality score 195 for an encode viewed on an additional viewing device based on the absolute quality score 195 of the encode viewed on the
5 base viewing device

[0045] Finally, at the end of the training phase, the perceptual quality calculator 160 receives the source models 155(1)-155(3) and the device equation 175(1). In alternate embodiments, the perceptual quality calculator 160 may receive any number of source models 155 and any number of device equations 175. During the prediction
10 phase, the perceptual quality calculator 160 receives target inputs. The “target inputs” include, without limitation, the target encode 185, the target source 187, the source resolution 180, and the target device type 190. In response to receiving the target inputs, as described in greater detail in conjunction with Figure 3, the perceptual quality calculator 160 generates the absolute quality score 195 for the
15 target encode 185 viewed on a target viewing device characterized by the target device type 190.

[0046] Until the perceptual quality calculator 160 receives a command to stop executing, as the perceptual quality calculator 160 receives new target inputs, the perceptual quality calculator 160 generates corresponding absolute quality scores
20 195. Advantageously, because the each of the absolute quality scores 195 reflects the source resolution 180 and the target device type 190, the absolute quality scores 195 correlate in a universal and absolute manner to subjective human visual experiences when viewing encodes on target viewing devices.

[0047] Advantageously, any number of the absolute quality scores 195 may be
25 aggregated to generate an aggregate quality score as part of evaluating the visual quality of streamed media content. For example, a dashboard application that monitors a streaming service on a daily basis can indicate the health of the streaming service based on the aggregated quality score. In general, unless a change occurs to the streaming service, the aggregated quality score remains stable. Consequently, a
30 deviation in the aggregated quality score between a particular day and the immediately proceeding day indicates an change to the streaming service (e.g., an issue or an improvement) that impacts the visual quality of the streamed media

content. By contrast, if a change to the streaming service is not intended to effect the visual quality of the streamed media content, then the aggregated quality score can provide confirmation that the visual quality of the streamed media content is unaffected. For instance, if a new encode for a smart phone saves bandwidth while
5 preserving visual quality, then the aggregated quality score remains stable while a time-weighted bitrate decreases.

[0048] In another example, an A/B test application compares a control group of streaming sessions with one or more treatment groups of streaming sessions. A treatment group differs from the control group in exactly one key parameter. The
10 aggregated quality score indicates whether the visual quality of the streamed media content is impacted by the difference in the key parameter. For instance, if the treatment group uses a new encode and the control group uses a current encode, then the A/B test application could compare the aggregated quality score for the control group and the aggregated quality score for the treatment group. The results of
15 the comparison indicates whether the new encode impacts the visual quality of streamed media content for the treatment groups.

[0049] Note that the techniques described herein are illustrative rather than restrictive, and may be altered without departing from the broader spirit and scope of the invention. Many modifications and variations on the functionality provided by the
20 subjective metric subsystem 140, the training subsystems 150, the conversion subsystems 170, and the perpetual quality calculator 160 will be apparent to those of ordinary skill in the art without departing from the scope and spirit of the described embodiments. For instance, in various embodiments, any number of the techniques or devices may be implemented while other techniques may be omitted or replaced in
25 any technically feasible fashion.

[0050] For instance, in some embodiments, the conversion subsystems 170 and the target device type 190 may be omitted from the absolute perceptual quality system 100. In such embodiments, the perceptual quality calculator 160 may be modified to set the absolute quality score 195 equal to the base quality score. In the same or
30 other embodiments, to increase the number of sources assigned to the metric engine 142, the metric engine 142 may include functionality to synthesize additional sources

having the assigned spatial resolution based on sources having higher spatial resolutions than the assigned spatial resolution.

[0051] It will be appreciated that the absolute perceptual quality system 100 shown herein is illustrative and that variations and modifications are possible. For example
5 the functionality provided by the subjective metric subsystem 140, the training subsystems 150, the conversion subsystems 170, and the perpetual quality calculator 160 as described herein may be integrated into or distributed across any number of software applications (including one) and any number of components of the absolute perceptual quality system 100. Further, the connection topology between the various
10 units in Figure 1 may be modified as desired.

[0052] Figure 2 is a more detailed illustration of the metric engine 142(2) of Figure 1, according to various embodiments of the present invention. As shown, the metric engine 142(2) includes, without limitation, a down-sampling engine 210, an encoding engine 220, a decoding engine 230, an up-sampling engine 240, and an objective
15 metric calculator 250. The metric engine 142(2) is associated with an assigned spatial resolution 242(2) of 720p. The assigned spatial resolution 242(2) corresponds to the spatial resolution of the sources included in the assigned source bucket 204(2) that the subjective metric subsystem 140 assigns to the metric engine 142(2).

[0053] As shown, the metric engine 142(2) receives the assigned source bucket
20 204(2) and a higher resolution source bucket 202(2). The assigned source bucket 204(2) includes sources having spatial resolutions of 720p, while the higher resolution bucket 202(2) includes sources having spatial resolutions of 1080p. The down-sampling engine 210 performs down-sampling operations on the higher resolution bucket 202(2) to synthesize additional 720p sources. The down-sampling engine 210
25 then adds the additional 720p sources to the assigned source bucket 204(2). In alternate embodiments, the metric engine 142(2) may omit the functionality to synthesize additional 720p sources, and the metric engine 142(2) does not receive the higher sources bucket 202(2).

[0054] Subsequently, the down-sampling engine 210 and the encoding engine 220
30 collaborate to generate encodes 225 derived from the sources included in the assigned source bucket 204(2). More precisely, the down-sampling engine 210 performs down-sampling operations on the sources included in the assigned source

bucket 204(2) to generate down-sampled sources that are included in four lower buckets 216. The lower bucket 216(1) includes down-sampled sources having spatial resolutions of 480p. The lower bucket 216(2) includes down-sample sources having spatial resolutions of 384p. The lower bucket 216(3) includes down-sampled sources
5 having spatial resolutions of 288p. The lower bucket 216(4) includes down-sample sources having spatial resolutions of 240p. The number and resolutions of the lower buckets 216 may be determined in any fashion that is consistent with the number and type of the encodes 225 desired.

[0055] The encoding engine 220 performs encoding operations on the sources
10 included in the assigned source bucket 204(2) and the down-sampled sources included in the lower buckets 216. More specifically for each source and down-sampled source, the encoding engine 220 generates one or more of the encodes 225 based on any number of encoding criteria and in any technically feasible fashion. For instance, in some embodiments, for each source and down-sampled source, the
15 encoding engine 220 generates three different encodes 225, where each of the encodes 225 corresponds to a different quality parameter. As shown, the encoding engine 220 generates the encodes 225(1)-225(E).

[0056] For each of the encodes 225, the decoding engine 230 performs decoding operations to generate a corresponding decode 235. As a result, the decoding
20 engine 230 generates the decodes 235(1)-235(E). As shown, the metric engine 142(2) provides the decodes 235 to a base viewing device 260 having a base device type of "smart television," an additional viewing device 270(1) having an additional device type of "smart phone," and an up-sampling engine 240. A smart phone is also referred to herein as a mobile device.

[0057] For each of the decodes 235(1)-235(E), humans 280 view the decodes 235(1)-
25 235(E) on the base viewing device 260 and generate associated raw absolute quality scores 195. For each of the decodes 235(1)-235(E), a mean opinion score calculator 290 processes the raw absolute quality scores 195 to generate a single absolute quality score 195. In alternate embodiments, the subjective metric subsystem 140
30 may process the raw absolute quality scores 195 to generate the quality score 195 in any technically feasible fashion. Accordingly, in some embodiments, the system 100 does not include the mean opinion score calculator 290.

[0058] The metric engine 140(2) then creates the subjective source bucket 125(2) that includes the absolute quality scores 195(1)-195(E) associated with the base viewing device 260. In addition, the subjective metric subsystem 140 adds the absolute quality scores 195(1)-195(E) associated with the base viewing device 260 to the base subjective bucket 172.

[0059] In a complementary fashion, for each of the decodes 235(1)-235(E), the humans 280 view the decodes 235(1)-235(E) on the additional viewing device 270(1) and generate associated raw absolute quality scores 195. For each of the decodes 235(1)-235(E), a mean opinion score calculation 290 processes the raw absolute quality scores 195 to generate a single absolute quality score 195 associated with the additional viewing device 270(1). The subjective metric subsystem 140 then adds the absolute quality scores 195(1)-195(E) associated with the additional viewing device 270(1) to the additional subjective bucket 174(1).

[0060] As persons skilled in the art will recognize, some of the encodes 225 generated by the metric engine 140(2) may be identical to some of the encodes 225 generated by one or more of the other metric engines 140(1) and 140(3). Notably, two encodes 225 that are identical have the same visual quality when decoded and viewed on a particular viewing device. Consequently, in some embodiments, the subjective metric subsystem 140 may reuse the absolute quality score 195 assigned by the humans 280 viewing one of the encodes 225 as the absolute quality score 195 for other, identical encodes 225. Accordingly, the humans 280 view only the unique encodes 225.

[0061] In alternate embodiments, the subjective metric subsystem 140 and/or the metric engines 144 may determine the absolute quality scores 195 assigned by the humans 280 in any technically feasible fashion. For example, for each unique encode 225 and each viewing device, the subjective metric subsystem 140 could retrieve raw values for any absolute metric from a database. Subsequently, the subjective metric subsystem 140 could process the raw values to generate the absolute quality score 195 associated with the viewing device.

[0062] The up-sampling engine 240 performs up-sampling operations 240 on the decodes 235(1)-235(E) based on the assigned spatial resolution 242(2) to generate re-constructed sources 245(1)-245(E) that match the spatial resolution of the sources

includes in the assigned source bucket 204(2). Subsequently, for each of the encodes 235(1)-235(E), the objective metric calculator 250 computes the corresponding objective value set 165 based on the associated source included in the assigned source bucket 204(2) and the associated re-constructed source 245.

- 5 **[0063]** The objective metric calculator 250 may compute the objective value set 165 in any technically feasible fashion. For instance, in some embodiments, the objective metric calculator 250 may compare the source to the re-constructed source 245 to determine one or more values included in the objective value set 195 that correspond to full-reference objective metrics. In some alternate embodiments, the objective
- 10 metric calculator 250 may implement one or more no-reference algorithms to compute one or more values included in the objective value set 195 that correspond to no-reference objective metrics. Because no-reference algorithms do not operate on the source, in some embodiments, the objective metric calculator 250 may not receive the assigned source bucket 204(2). Finally, the objective metric calculator 250 creates
- 15 the objective source bucket 135(2) that includes the objective value sets 165(1)-165(E).

- [0064]** As persons skilled in the art will recognize, each objective metric may represent any number of deterministic video characteristics and exhibits both strengths and weaknesses. For example, a visual information fidelity (VIF) is
- 20 relatively good at detecting slight blurring artifacts, but is relatively poor at detecting blocking artifacts. In some embodiments, to leverage the strengths and mitigate the weaknesses, the objective metric calculator 250 may be configured to determine the set of objective metrics 145 that, together, provide valuable insight into the visual quality across a wide range of the encodes 225. The selection of the set of objective
- 25 metrics may be made in any technically feasible fashion to address any number of anticipated artifacts. For instance, in some embodiments, the set of objective metrics may be empirically selected to assess degradation caused by compression (i.e., blockiness) and scaling (i.e., blurriness).

- [0065]** More precisely, in some embodiments, the set of objective metric include an
- 30 additive impairment measure (ADM), a visual information fidelity (VIF), a blind/referenceless image spatial quality evaluator (BRISQUE), and a temporal information (TI). Each of these objective metrics may be associated with any number

of values, also known as “features.” As part of generating the objective value set 165, the objective metric calculator 250 may include any number of features for each objective metric and discard the remainder. For example, in some embodiments, the objective value set 165 specifies values for one ADM feature, four VIF features,
 5 twelve BRISQUE features, and one TI feature.

[0066] Many modifications and variations on the functionality provided by the subjective metric subsystem 140 and the metric engine 142(2) will be apparent to those of ordinary skill in the art without departing from the scope and spirit of the described embodiments. In particular, modifications of the functionality described in
 10 Figure 2 to reflect different assigned spatial resolutions 242 (e.g., for metric engines 142(1) and 142(3)) will be apparent to those of ordinary skill in the art.

Computing Absolute Quality Scores

[0067] Figure 3 is a more detailed illustration of the perceptual quality calculator 160 of Figure 1, according to various embodiments of the present invention. As shown,
 15 the perceptual quality calculator 160 includes, without limitation, a model selector 310, a base calculator 320, and a viewing device adjuster 330. For explanatory purposes only, the base device type is smart television. Further, at the end of the training phase, the perceptual quality calculator 160 receives the source model 155(1), the source model 155(2), the source model 155(3), the device equation 175(1), and the
 20 device equation 175(2).

[0068] The source model 155(1) is associated with the spatial resolution of 480p. The source model 155(2) is associated with the spatial resolution of 720p. The source model 155(3) is associated with the spatial resolution of 1080p. The device equation 175(1) is associated with the additional device type of smart phone. More specifically,
 25 the device equation 175(1) expresses the absolute quality score 195 associated with a device type of smart phone as a function of a base quality score 365. The base quality score 365 is the absolute quality score 195 associated with the base device type of smart television. For example, the device equation 175(1) could be the following polynomial:

$$\begin{aligned}
 & \text{(absolute quality score 195) =} & (1) \\
 & 0.08492679 * (\text{base quality score 365}) + \\
 & 0.00879799 * (\text{base quality score 365})^2
 \end{aligned}$$

[0069] The device equation 175(2) is associated with the additional device type of laptop. To reflect that the additional device type of laptop is to be considered visually equivalent to the base device type, the associated device equation 175(2) may explicitly set the absolute quality score 195 equal to the base equality score 365. For
 5 example, the device equation 175(2) could be:

$$(\text{absolute quality score } 195) = (\text{base quality score } 365) \quad (2)$$

[0070] In alternate embodiments, the perceptual quality calculator 160 and/or the viewing device adjuster 330 may determine that an additional device type is to be considered visually equivalent to the base device type in any technically feasible. In
 10 various embodiments, the base device type may include a list of device types. In alternate embodiments, the base device type may specify any number and type of characteristics associated with any number of viewing devices. For example, the base device type could specify a lists of resolutions associated with viewing devices.

[0071] During the prediction phase, the perceptual quality calculator 160 receives the
 15 source resolution 180, the target source 187, the target encode 185, and the target device type 190. In some embodiments, the perceptual quality calculator 160 may not receive the target device type 190. In such embodiments, the perceptual quality calculator 160 sets the target device type 190 equal to the base device type.

[0072] As shown, the model selector 310 selects one of the source models 155 as the
 20 target model 355 based on the source resolution 180. More precisely, the model selector 310 performs one or more comparison operations between the source resolution 180 and the different spatial resolutions associated with the source models 155. If the source resolution 180 matches the spatial resolution associated with one of the source models 155, then the model selector 310 selects the source model 155
 25 as the target model 355. For example, if the source resolution 180 was equal to 1080p, then the model selector 310 would select the source model 155(3) as the target model 355.

[0073] If, however, the source resolution 180 does not match the spatial resolution associated with one of the source models 155, then the model selector 310 selects
 30 the source model 155 associated with the highest spatial resolution that is not greater than the source resolution 180. For example, if the source resolution 180 was equal

to 600p, then the model selector 310 would select the source model 155(1) as the target model 355. In alternate embodiments, the model selector 310 may determine the target model 355 in any technically feasible fashion that reflects the source resolution 180.

- 5 **[0074]** Subsequently, the base calculator 320 computes a base quality score 365 based on the target model 355, the target encode 185, and the target source 187. First, the base calculator 320 computes the objective value set 165 associated with the target encode 187. The base calculator 320 may compute the objective value set 165 in any technically feasible fashion. For instance, in some embodiments, the base
10 calculator 320 performs decoding operations to generate a target decode. The base calculator 320 performs up-sampling operations on the target decode to generate a re-constructed target source having a spatial resolution equal to the source resolution 180. Subsequently, the base calculator 320 performs comparison operations between the target source 187 and the re-constructed target source to determine one
15 or more values included in the objective value set 195 that correspond to full-reference objective metrics.

- [0075]** In some alternate embodiments, the base calculator 320 may implement one or more no-reference algorithms to compute one or more values included in the objective value set 195 that correspond to no-reference objective metrics. Because
20 no-reference algorithms do not operate on the source, in some embodiments, the perceptual quality calculator 160 may not receive the target source 187. Finally, the base calculator 320 computes the base quality score 365 based on the target model 355 and the objective metric set 165.

- [0076]** The viewing device adjuster 330 compares the target device type 190 to the
25 base device type. If the target device type 190 is equal to the base device type, then the viewing device adjuster 330 sets the absolute quality score 195 equal to the base quality score 365. For example, if the target device type 190(1) is smart television, then the viewing device adjuster 330 sets the absolute quality score 195 equal to the base quality score 365.

- 30 **[0077]** If, however, the target device type 190 is not equal to the base device type, then the viewing device adjuster 330 selects one of the device equations 175 as a target device equation 340 based on the target device type 190. More precisely, the

viewing device adjuster 330 performs comparison operations between the target device type 190 and the additional device types associated with the device equations 175. The target device adjuster 330 then selects the device equation 175 associated with the additional device type that matches the target device type 190 as the target device equation 340. For example, if the target device type 190(1) is smart phone, then the viewing device adjuster 330 selects the device equation 175(1) as the target device equation 340. The viewing device adjuster 330 then computes the absolute quality score 195 associated with the target device type based on the target device equation 340 and the base quality score 365.

10 **[0078]** Figure 4 is a flow diagram of method steps for computing an absolute quality score for encoded video content viewed on a base viewing device, according to various embodiments of the present invention. Although the method steps are described with reference to the systems of Figures 1-3, 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.

[0079] As shown, a method 400 begins at step 402, where the subjective metric subsystem 140 receives sources and assigns each of the sources to one of the assigned source buckets 204 based on the spatial resolution of the source. More specifically, the subjective metric subsystem 140 assigns the source to the assigned source bucket 204(i) corresponding to the assigned spatial resolution 242(i) that equals the spatial resolution of the source. As described previously herein, each of the assigned source buckets 204(i) is also associated with a different metric engine 142(i).

25 **[0080]** At step 404, for each of the assigned spatial resolutions 242(i), the down-sampling engine 210 performs down-sampling operations to generate additional sources from sources included in the higher source resolution bucket 202(i). The down-sampling engine 210 then adds the additional sources to the corresponding assigned source bucket 204(i).

30 **[0081]** At step 406, for each of the assigned spatial resolutions 242(i), the down-sampling engine 210 and the encoding engine 220 collaborate to generate one or more encodes 225 for each of the sources included in the corresponding assigned

source bucket 204(i). Subsequently, as part of step 406, the decoding engine 230 performs decoding operations on the encodes 225 to generate the decodes 235.

[0082] At step 408, for each of the assigned spatial resolutions 242(i), the up-sampling engine 240 performs up-sampling operations on the decodes 235 to
5 generate re-constructed sources 245. Notably, the re-constructed sources 245 and the sources included in the corresponding assigned source bucket 204(i) have the assigned spatial resolution 242(i). At step 410, for each of the assigned spatial resolutions 242(i), the objective metric calculator 250 computes the objective value sets 165 included in the corresponding objective source bucket 135(i) based on the
10 re-constructed sources 245 and the corresponding assigned source bucket 204(i).

[0083] At step 412, for each of the assigned spatial resolutions 242(i), one of the training subsystems 150 generates the associated source model 155(i) via a machine learning algorithm trained on the objective source bucket 135(i) and the corresponding subjective source bucket 125(i). Importantly, each objective value set
15 165 included in the objective source bucket 135(i) is associated with a particular encode 225, and the corresponding subjective source bucket 125(i) includes the absolute quality score 195 for the encode 225. The absolute quality score 195 for a particular encode 225 is generated based on human video-viewings of the encode 225 on the base viewing device 260. Accordingly, during step 412, each of the
20 source models 155(i) is trained to associate the objective value set 165 and the absolute quality score 195 for encodes 225 derived from sources having the assigned spatial resolution 242(i).

[0084] At step 414, the perceptual quality calculator 160 receives the source resolution 180, the target source 187, and the target encode 185. Because the
25 perceptual quality calculator 160 does not receive the target device type 190, the perceptual quality calculator 160 sets the target device type 190 equal to the base device type. At step 416, the model selector 310 sets the target model 355 equal to the source model 155 based on the source resolution 180.

[0085] At step 418, the base calculator 320 computes the objective value set 165
30 associated with the target encode 185 based on the target encode 185 and the target source 187. At step 420, the base calculator 320 computes the base quality score 365 based on the target model 355 and the objective value set 165 associated with

the target encode 185. Because the target device type 190 is equal to the base device type, the viewing device adjuster 330 sets the absolute quality score 195 equal to the base quality score 365. Subsequently, any number of the absolute quality scores 195 may be aggregated to generate an aggregate quality score as part of evaluating the visual quality of streamed media content. The method 400 then terminates.

[0086] Figure 5 is a flow diagram of method steps for computing an absolute quality score for encoded video content viewed on a target viewing device, according to various embodiments of the present invention. Although the method steps are described with reference to the systems of Figures 1-3, 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.

[0087] As shown, a method 500 begins at step 502, where the perceptual quality modeling subsystem 100 receives the base subjective bucket 172 and the additional subjective buckets 174. Each of the additional subjective buckets 174 corresponds to an different additional viewing device 270. At step 504, for each additional viewing device 270(i), the conversion subsystem 170 generates a plot based on the base subjective bucket 172 and the additional subjective bucket 174(i) associated with the additional viewing device 270.

[0088] Each point included in the plot is associated with a different one of the encodes 225. The horizontal coordinate of a particular point represents the absolute quality score 195 for the associated encode 225 that is included in the base subjective bucket 172. The vertical coordinate of the point represents the absolute quality score 195 for the associated encode 225 that is included in the additional subjective bucket 174(i). At step 506, for each additional viewing device 270(i), the conversion subsystem 170 performs one or more curve-fitting operations to generate the associated device equation 175(i).

[0089] At step 508, the perceptual quality calculator 160 receives the source resolution 180, the target source 187, the target encode 185, and the target device type 190. At step 510, the perceptual quality calculator 160 computes the base quality score 365 associated with the target encode 185. The perceptual quality calculator 160 may compute the base quality score 365 in any technically feasible

fashion. For example, the perceptual quality calculator 160 could implement the method steps 416-420 of method 400 described in conjunction with Figure 4.

5 [0090] At step 512, the viewing device adjuster 330 determines whether the target device type 190 is equal to the base device type. If, at step 512, the viewing device adjuster 330 determines that the target device type 190 is not equal to the base device type, then the method 500 proceeds to step 514. At step 514, the viewing device adjuster 300 sets the absolute viewing score 195 equal to the base quality score 365, and the method 500 terminates.

10 [0091] If, however, at step 512, the viewing device adjuster 330 determines that the target device type 190 is equal to the base device type, then the method 500 proceeds directly to step 516. At step 516, the viewing device adjuster 330 sets the target device equation 340 equal to one of the device equations 175 based on the target device type 190. At step 518, the viewing device adjuster 330 computes the absolute viewing score 195 based on the target device equation 340 and the base
15 quality score 365. Subsequently, any number of the absolute quality scores 195 may be aggregated to generate an aggregate quality score as part of evaluating the visual quality of streamed media content. The method 500 then terminates.

[0092] In sum, the disclosed techniques may be used to efficiently and reliably predict an absolute quality score for encodes derived from sources. Initially, a subjective
20 metric subsystem divides sources into assigned source buckets based on the spatial resolutions of the sources. Each assigned source bucket is associated with an assigned source resolution. For each assigned source bucket, the subjective metric subsystem derives encodes from each source and generates an associated objective source bucket that includes an objective value set for each of the encodes.

25 [0093] Subsequently, for each assigned source bucket, a training subsystem trains a corresponding source model based on the associated objective bucket and a corresponding subjective source bucket. For each encode associated with the assigned bucket, the subjective source bucket includes a “base” absolute quality score that is assigned based on human video-viewings of the encode on a base
30 viewing device. The base viewing device is associated with a base device type. In this fashion, each source model is trained to associate an objective value set for an

encode associated with a particular source resolution and a base quality score for the encode.

[0094] For each additional device type, a conversion subsystem generates a corresponding device equation based on a base subjective bucket and a device
5 subjective bucket associated with the additional device type. For each encode generated by the metric engine, the base subjective bucket includes the base quality score. For each encode generated by the subjective metric subsystem, the device subjective bucket associated with a given additional device type includes the absolute
10 quality score associated with the additional device type. The absolute quality score associated with an additional device type is assigned based on human video-viewings of the encode on an additional viewing device that is associated with the associated additional device type.

[0095] To determine the absolute quality score for a target encode, a perceptual quality calculator sets a target model equal to the source model associated with the
15 spatial resolution of the source from which the target encode was derived. The perceptual quality calculator then computes a base quality score associated with the target encode based on the target model and a set of objective metrics for the target encode. If the target device type associated with the viewing device on which the target encode is to be viewed is equal to the base device type, then the perceptual
20 quality calculator sets the absolute quality score associated with the encode equal to the base quality score associated with the encode.

[0096] If, however, the target device type does not equal the base device type, then the perceptual quality calculator sets a target device equation equal to the device equation associated with the target device type. The perceptual quality calculator
25 then computes the absolute quality score for the target encode based on the target device equation and the base quality score for the target encode.

[0097] Advantageously, the perceptual quality calculator accurately predicts the visual quality of encodes derived from sources having different spatial resolutions as
perceived when viewed on different viewing devices. More specifically, training
30 different models based on sources having different spatial resolutions enables the perceptual quality calculator to accurately compute base quality scores for encodes derived from sources having a wide range of spatial resolutions. In a complementary

fashion, generating device equations for different viewing devices enables the perceptual quality calculator to accurately adjust base quality scores to reflect the target viewing device. Consequently, the absolute quality scores allow meaningful comparison and effective optimization operations across encodes derived from a wide
5 range of sources and viewed on viewing devices with diverse viewing characteristics.

[0098] 1. In some embodiments, a method comprises selecting a first model based on a first spatial resolution of first video content, wherein the first model is included in a plurality of models and associates a set of objective values for a set of objective quality metrics with an absolute quality score; computing a first set of values for the
10 set of objective quality metrics based on first encoded video content derived from the first video content; and computing a first absolute quality score for the first encoded video content based on the first model and the first set of values, wherein the first absolute quality score is aggregated with one or more other absolute quality scores in evaluating the visual quality of streamed media content.

15 **[0099]** 2. The method of clause 1, wherein computing the first set of values comprises decoding the first encoded video content to generate a first decoded video content; up-sampling the first decoded video content to generate first re-constructed video content having the first spatial resolution; and computing a value for a first objective quality metric included in the set of objective quality metrics based on the
20 first re-constructed video content and the first video content.

[0100] 3. The method of clauses 1 or 2, wherein the first set of values includes values for at least one of temporal information, a detail loss measure, and visual information fidelity.

25 **[0101]** 4. The method of any of clauses 1-3, wherein each model included in the plurality of models is associated with a different spatial resolution included in a plurality of spatial resolutions, and selecting the first model comprises determining a highest spatial resolution included in the plurality of resolutions that is not greater than the first spatial resolution; and identifying a model included in the plurality of models that is associated with the highest resolution.

30 **[0102]** 5. The method of any of clauses 1-4, further comprising, prior to selecting the first model encoding training video content having the first spatial resolution to

generate encoded training video content; generating training data based on the encoded training video content; and generating the first model via a machine learning algorithm trained using the training data.

5 [0103] 6. The method of any of clauses 1-5, wherein generating the training data comprises computing a second set of values for the set of objective quality metrics based on the encoded training video content; and associating the second set of values with a second absolute quality score for the encoded training video content.

10 [0104] 7. The method of any of clauses 1-6, further comprising generating the second absolute quality score based on one or more human-observed visual quality scores for re-constructed video content derived from the encoded training video content.

[0105] 8. The method of any of clauses 1-7, wherein the first spatial resolution comprises 480p, 720p, 1080p, or 4K.

15 [0106] 9. The method of any of clauses 1-8, further comprising selecting a second model based on a second spatial resolution of second video content, wherein the second model is included in the plurality of models; computing a second set of values for the set of objective quality metrics based on second encoded video content derived from the second video content; and computing the second absolute quality score for the second encoded video content based on the second model and the
20 second set of values, wherein the second absolute quality score is aggregated with the first absolute quality score in evaluating the visual quality of streamed media content.

25 [0107] 10. In some embodiments, a computer-readable storage medium includes instructions that, when executed by a processor, cause the processor to perform the steps of selecting a first model based on a first spatial resolution of first video content, wherein the first model is included in a plurality of models; computing a first set of values for a set of objective quality metrics based on the first video content and first encoded video content derived from the first video content, wherein the set of objective quality metrics represent a plurality of deterministic video characteristics;
30 and computing a first absolute quality score for the first encoded video content based on the first model and the first set of values, wherein the first absolute quality score is

aggregated with one or more other absolute quality scores in evaluating the visual quality of streamed media content.

5 [0108] 11. The computer-readable storage medium of clause 10, wherein computing the first set of values comprises decoding the first encoded video content to generate a first decoded video content; up-sampling the first decoded video content to generate first re-constructed video content having the first spatial resolution; and computing a value for a first objective quality metric included in the set of objective quality metrics based on the first re-constructed video content and the first video content.

10 [0109] 12. The computer-readable storage medium of clauses 10 or 11, wherein computing the value for the first objective quality metric comprises performing one or more comparison operations between the first re-constructed video content and the first video content.

15 [0110] 13. The computer-readable storage medium of any of clauses 10-12, wherein each model included in the plurality of models is associated with a different spatial resolution included in a plurality of spatial resolutions, and selecting the first model comprises determining a highest spatial resolution included in the plurality of resolutions that is not greater than the first spatial resolution; and identifying a model included in the plurality of models that is associated with the highest resolution.

20 [0111] 14. The computer-readable storage medium of any of clauses 10-13, further comprising, prior to selecting the first model encoding training video content having the first spatial resolution to generate encoded training video content; generating training data based on the encoded training video content; and generating the first model via a machine learning algorithm trained using the training data.

25 [0112] 15. The computer-readable storage medium of any of clauses 10-14, further comprising, prior to encoding the training video content, deriving the training video content from base video content having a second spatial resolution that exceeds the first spatial resolution.

30 [0113] 16. The computer-readable storage medium of any of clauses 10-15, wherein the machine learning algorithm comprises a support vector machine algorithm, an artificial neural network algorithm, or a random forest algorithm.

5 **[0114]** 17. The computer-readable storage medium of any of clauses 10-16, wherein generating the training data comprises computing a second set of values for the set of objective quality metrics based on the encoded training video content; and associating the second set of values with a second absolute quality score for the encoded training video content.

10 **[0115]** 18. The computer-readable storage medium of any of clauses 10-17, further comprising selecting a second model based on a second spatial resolution of second video content, wherein the second model is included in the plurality of models; computing a second set of values for the set of objective quality metrics based on the second video content and second encoded video content derived from the second video content; and computing the second absolute quality score for the second encoded video content based on the second model and the second set of values, wherein the second absolute quality score is aggregated with the first absolute quality score in evaluating the visual quality of streamed media content.

15 **[0116]** 19. In some embodiments, a system comprises a memory storing a perceptual quality application; and a processor coupled to the memory, wherein when executed by the processor, the perceptual quality application causes the processor to select a first model based on a first spatial resolution of first video content, wherein the first model is included in a plurality of models; generate first re-constructed video content
20 based on first encoded video content derived from the first video content; and compute a first absolute quality score for the first encoded video content based on the first model and the first re-constructed video content, wherein the first absolute quality score is aggregated with one or more other absolute quality scores in evaluating the visual quality of streamed media content.

25 **[0117]** 20. The system of clause 19, wherein the perceptual quality application causes the processor to compute the first absolute quality score by computing a first set of values for a set of objective quality metrics based on the first re-constructed video content; and mapping the first set of values to the first absolute quality score via the first model.

30 **[0118]** Any and all combinations of any of the claim elements recited in any of the claims and/or any elements described in this application, in any fashion, fall within the contemplated scope of the present invention and protection.

5 [0119] The descriptions of the various embodiments have been presented for purposes of illustration, but are not intended to be exhaustive or limited to the 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 embodiments.

10 [0120] Aspects of the present embodiments may be embodied as a system, 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 “module” or “system.” Furthermore, aspects of the 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.

15 [0121] 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 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 program for use by or in connection with an instruction execution system, apparatus, or device.

30 [0122] 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 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 machine. The instructions, when executed 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 gate arrays.
- 10 **[0123]** 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 represent a module, segment, or portion of code, which comprises one or more
- 15 executable instructions for implementing the specified logical function(s). It should 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
- 20 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 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.
- 25 **[0124]** While the preceding is directed to embodiments of the present disclosure, 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.

WHAT IS CLAIMED IS:

1. A computer-implemented method for computing an absolute quality score for encoded video content viewed on a base viewing device, the method being performed by a perceptual quality modeling system comprising a processor and memory, the method comprising the steps of:

selecting, by the processor, a first model based on a first spatial resolution of first video content, wherein the first model is included in a plurality of models and associates a set of objective values for a set of objective quality metrics with an absolute quality score, and each model included in the plurality of models was trained using different source video content having a same spatial resolution;

computing, by the processor, a first set of values for the set of objective quality metrics based on first encoded video content derived from the first video content, wherein each objective quality metric corresponds to one or more deterministic video characteristics; and

computing, by the processor, a first absolute quality score for the first encoded video content by:

computing a base quality score based on the first model and the first set of values,

comparing a type of a target viewing device with a type of the base viewing device,

selecting a device equation based on the target viewing device type when the target viewing device type does not equal the base viewing device type; and

computing the first absolute quality score based on the target viewing device equation and the base quality score.

2. The computer-implemented method of claim 1, wherein computing the first set of values comprises:

decoding, by the processor, the first encoded video content to generate a

first decoded video content;

up-sampling, by the processor, the first decoded video content to generate first re-constructed video content having the first spatial resolution;
and

computing, by the processor, a value for a first objective quality metric included in the set of objective quality metrics based on the first re-constructed video content and the first video content.

3. The computer-implemented method of claim 2, wherein computing the value for the first objective quality metric comprises performing one or more comparison operations between the first re-constructed video content and the first video content.

4. The computer-implemented method of claim 1, wherein the first set of values includes values for at least one of temporal information, a detail loss measure, and visual information fidelity.

5. The computer-implemented method of claim 1, wherein each model included in the plurality of models is associated with a different spatial resolution included in a plurality of spatial resolutions, and selecting the first model comprises:

determining a highest spatial resolution included in the plurality of resolutions that is not greater than the first spatial resolution; and

identifying a model included in the plurality of models that is associated with the highest resolution.

6. The computer-implemented method of claim 1, further comprising, prior to selecting the first model:

encoding training video content having the first spatial resolution to generate encoded training video content;

generating training data based on the encoded training video content; and

generating the first model via a machine learning algorithm trained using the training data.

7. The computer-implemented method of claim 6, wherein generating the training data comprises:

computing a second set of values for the set of objective quality metrics based on the encoded training video content; and

associating the second set of values with a second absolute quality score for the encoded training video content.

8. The computer-implemented method of claim 7, further comprising generating the second absolute quality score based on one or more human-observed visual quality scores for re-constructed video content derived from the encoded training video content.

9. The computer-implemented method of claim 6, further comprising, prior to encoding the training video content, deriving the training video content from base video content having a second spatial resolution that exceeds the first spatial resolution.

10. The computer implemented method of claim 6, wherein the machine learning algorithm comprises a support vector machine algorithm, an artificial neural network algorithm, or a random forest algorithm.

11. The computer-implemented method of claim 1, wherein the first spatial resolution comprises 480p, 720p, 1080p, or 4K.

12. The computer-implemented method of claim 1, further comprising:
selecting a second model based on a second spatial resolution of
second

video content, wherein the second model is included in the plurality of models;

computing a second set of values for the set of objective quality metrics based on second encoded video content derived from the second video content; and

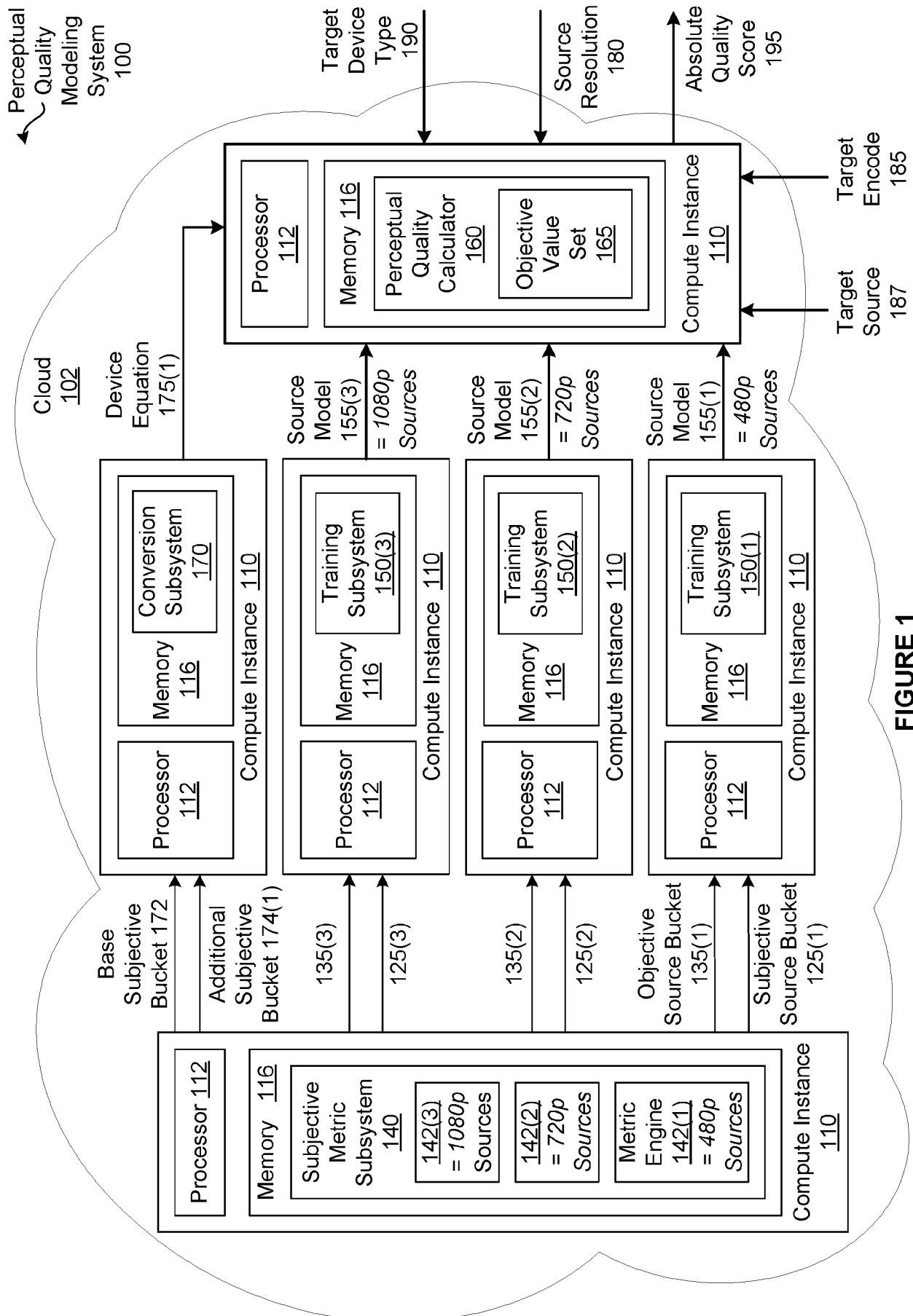
computing the second absolute quality score for the second encoded video content based on the second model and the second set of values,

wherein the second absolute quality score is aggregated with the first absolute quality score in evaluating the visual quality of streamed media content.

13. A non-transitory computer-readable storage medium including instructions that, when executed by a processor cause the processor to perform the steps of any one of claims 1 to 12.

14. A system comprising:

a memory storing a perceptual quality application and a processor coupled to the memory, wherein when executed by the processor, the perceptual quality application causes the processor to perform the method of any one of claims 1 to 12.



2/5

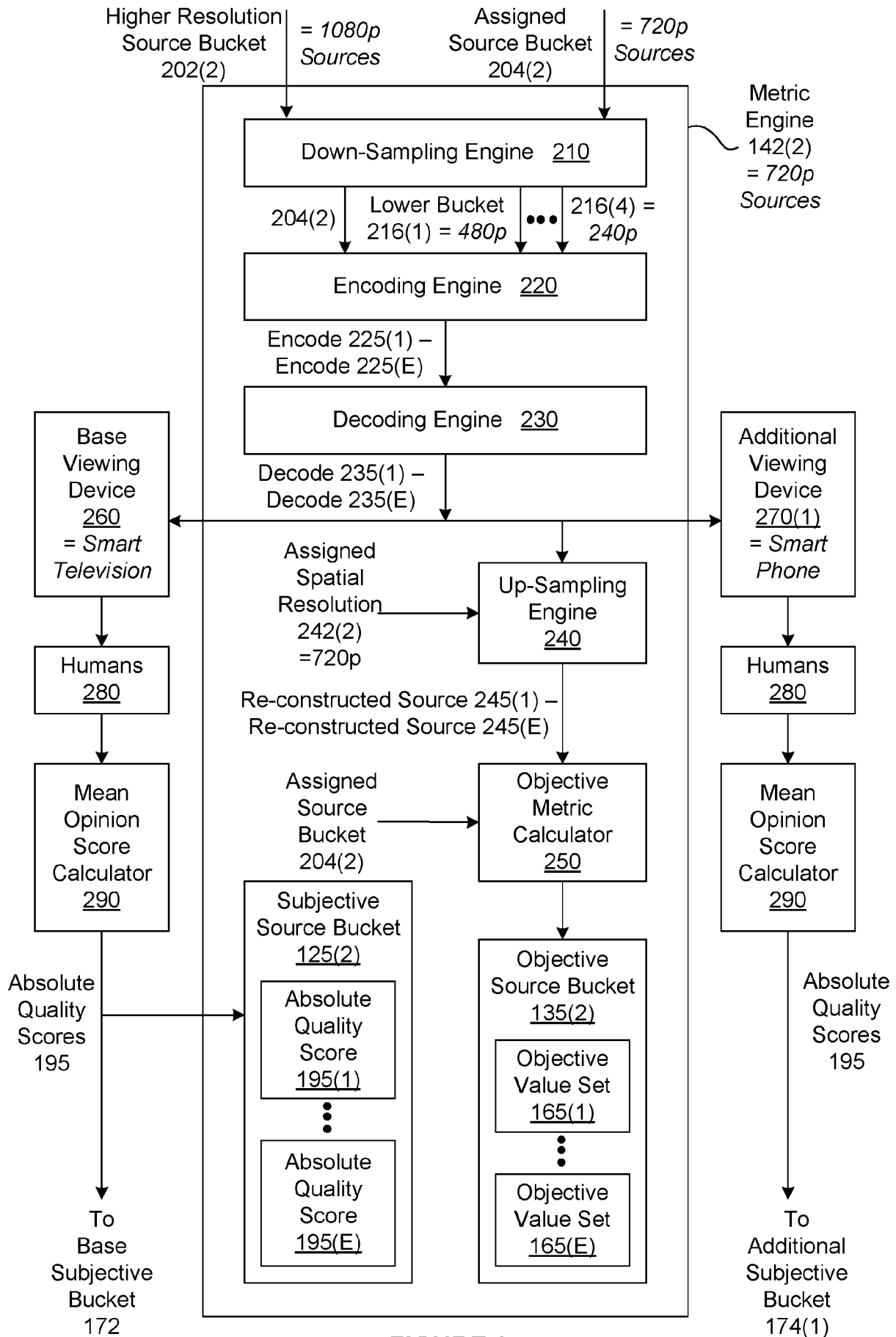


FIGURE 2

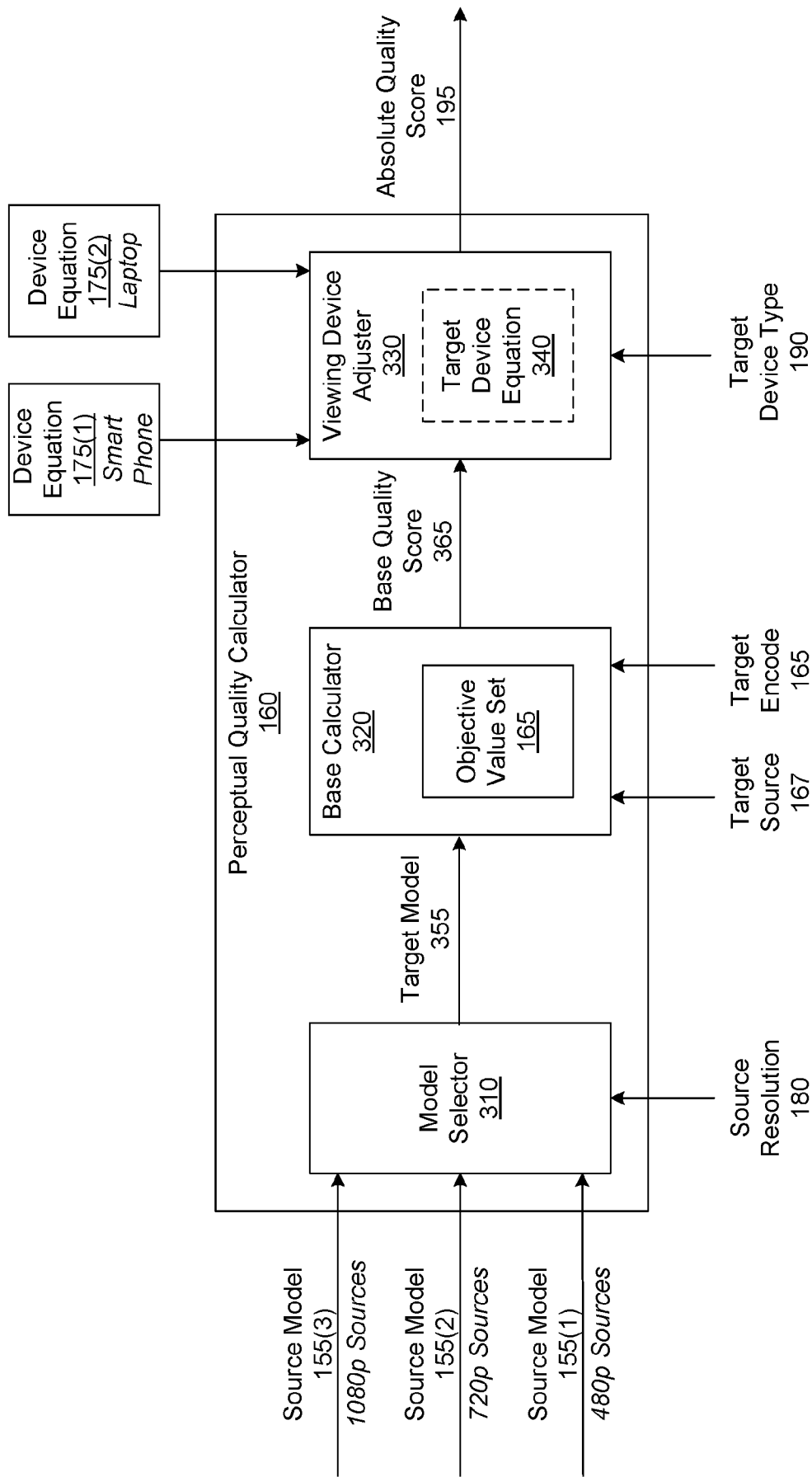


FIGURE 3

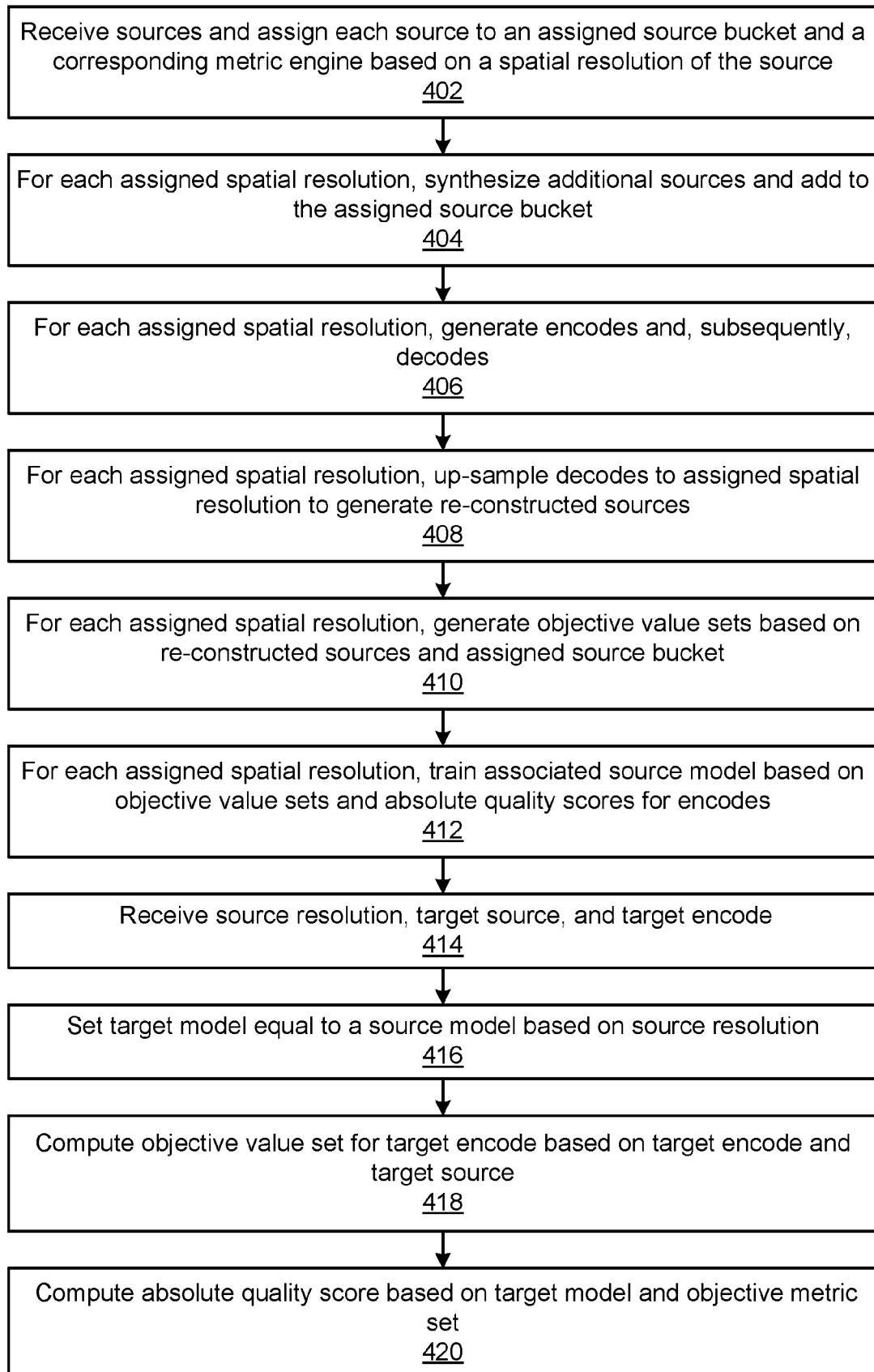


FIGURE 4

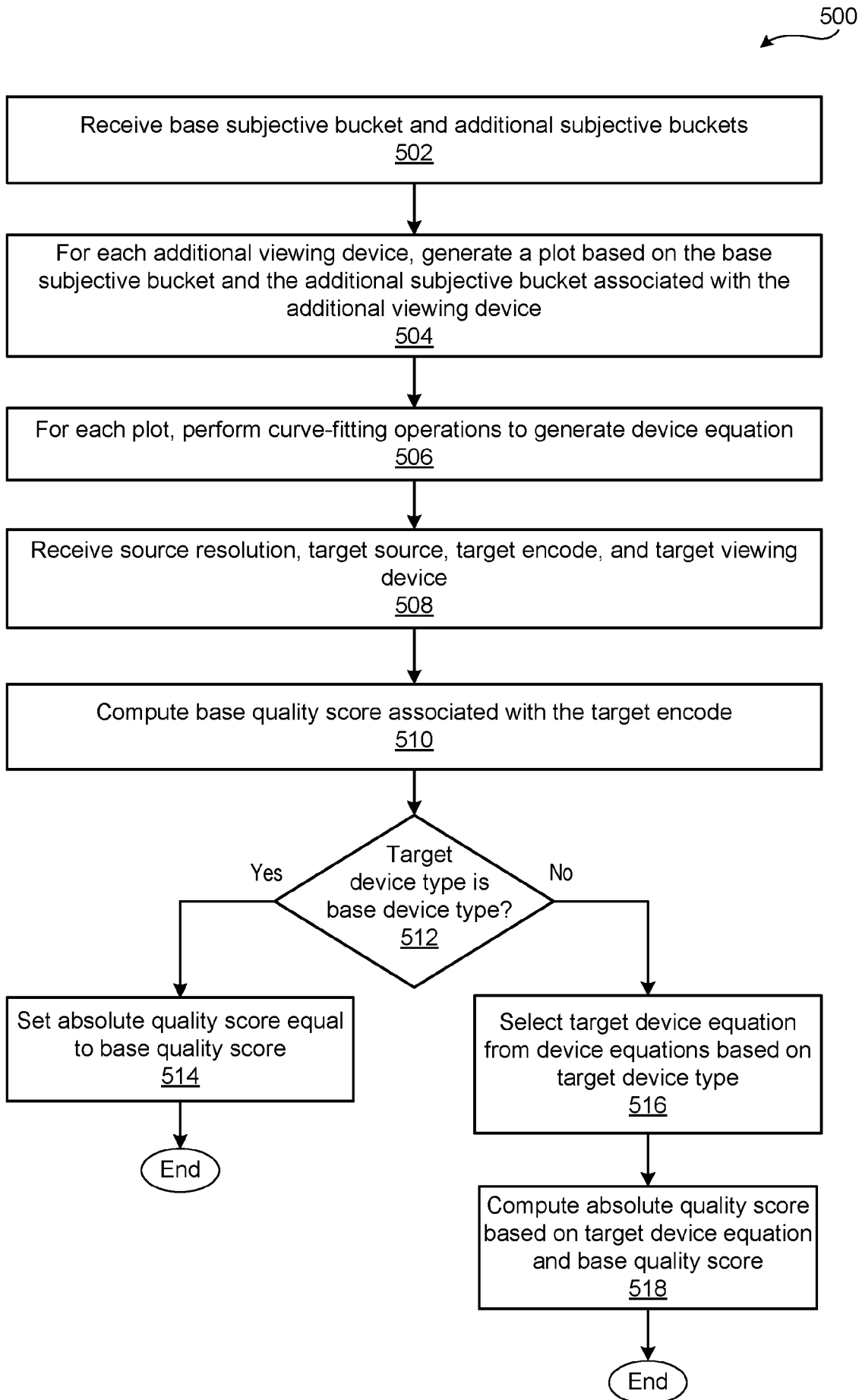


FIGURE 5

