



- (51) **International Patent Classification:**
H04N 17/00 (2006.01) *H04N 19/154* (2014.01)
- (21) **International Application Number:**
PCT/IB2015/054313
- (22) **International Filing Date:**
8 June 2015 (08.06.2015)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
14/336,246 21 July 2014 (21.07.2014) US
- (71) **Applicant:** CISCO TECHNOLOGY, INC. [US/US]; 170 West Tasman Drive, San Jose, California 95134-1706 (US).
- (72) **Inventors:** LI, Zhi; 2700 Del Medio Ct., Apt 315, Mountain View, California 94040 (US). GAHM, Josh; 55 Washington Park, Newtonville, Massachusetts 02460 (US).
- (74) **Agents:** KATZ, Samuel M. et al.; Cisco Legal (Patents), Cisc Technology, Inc., 170 West Tasman Drive, San Jose, CA 95134-1706 (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, LI, 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: RESOLUTION ROBUST VIDEO QUALITY METRIC

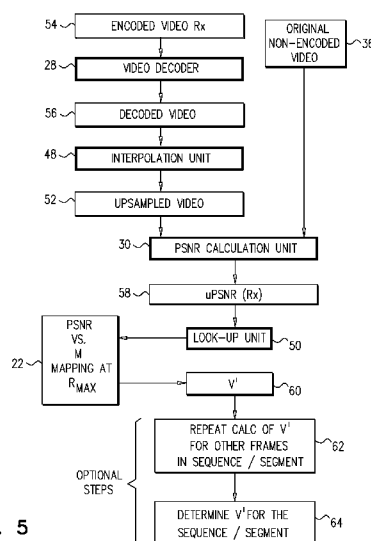


FIG. 5

(57) **Abstract:** In one embodiment, a video processing system includes a processor to read a mapping between a peak signal to noise ratio, PSNR, and a video quality metric, M , for an encoded version of a first video portion encoded at full resolution R_{max} , read an up-sampled peak signal to noise ratio, $uPSNR(Rx)$ of an up-sampled decoded video portion deriving from the first video portion being encoded at a resolution R_x , which is less than the full resolution R_{max} , yielding an encoded video portion, the encoded video portion being decoded and then up-sampled to the full resolution R_{max} ; and lookup a value, V' , that maps to $uPSNR(Rx)$ in the mapping, V' representing a resolution normalized M value of the encoded video portion encoded at the resolution R_x . Related apparatus and methods are also described.

RESOLUTION ROBUST VIDEO QUALITY METRIC

TECHNICAL FIELD

[0001] The present disclosure relates generally to video quality metrics.

5

BACKGROUND

[0002] In video streaming applications, it may be important to determine a measure of the quality of the video delivered to end-users. Such a metric may help gauge user satisfaction. Additionally, when combined with other information (such as video bitrate), it may be used to optimize users' viewing experience given the same bandwidth resource, or maintain the same level of viewing experience while packing more users within the same bandwidth.

[0003] An ideal video quality metric typically has a high degree of accuracy, in that the generated quality score closely matches the quality perceived by end users. Additionally, an ideal video quality metric is typically easy to compute, so that it may even be applied in real-time applications such as live streaming.

BRIEF DESCRIPTION OF THE DRAWINGS

[0004] The present disclosure will be understood and appreciated more fully from the following detailed description, taken in conjunction with the drawings in which:

5 [0005] Fig. 1 is a block diagram view of a video processing system constructed and operative in accordance with an embodiment of the present invention;

[0006] Fig. 2 is a chart showing a mapping between uPSNR and M at various resolutions for use in the system of Fig. 1;

10 [0007] Fig. 3 is a flow chart type view showing a method of determining the mapping between PSNR and M at full resolution for use in the system of Fig. 1;

[0008] Fig. 4 is a view of the chart of Fig. 2 showing looking up uPSNR(Rx) value in the mapping to determine a resolution normalized M value;

15 [0009] Fig. 5 is a flow chart type view showing a method to determine the resolution normalized M value of Fig. 4;

[0010] Fig. 6 is a block diagram view of a client device constructed and operative in accordance with an embodiment of the present invention;

20 [0011] Fig. 7 is a flow chart showing a method of selecting bitrate segments in the client device of Fig. 6;

[0012] Fig. 8 is a block diagram view of a device constructed and operative in accordance with an embodiment of the present invention; and

25 [0013] Fig. 9 is a flow chart showing a method of calculating a video quality metric for the segments selected by the client device of Fig. 6.

DESCRIPTION OF EXAMPLE EMBODIMENTS

OVERVIEW

[0014] There is provided in accordance with an embodiment of the present invention, a video processing system including a processor, and a memory to store data used by the processor, wherein the processor is operative to read, from the memory, a mapping between a peak signal to noise ratio, PSNR, and a video quality metric, M , for an encoded version of a first video portion encoded at full resolution R_{max} , read, from the memory, an up-sampled peak signal to noise ratio, $uPSNR(R_x)$, of an up-sampled decoded video portion, the up-sampled decoded video portion deriving from the first video portion being encoded at a resolution R_x , which is less than the full resolution R_{max} , yielding an encoded video portion, the encoded video portion being decoded and then up-sampled to the full resolution R_{max} , and lookup a value, V' , that maps to $uPSNR(R_x)$ in the mapping V' representing a resolution normalized M value of the encoded video portion encoded at the resolution R_x .

DESCRIPTION CONTINUED

[0015] The term “encoded” is used throughout the present specification and claims, in all of its grammatical forms, to refer to any type of data stream encoding including, for example and without limiting the scope of the definition, well known types of encoding such as, but not limited to, MPEG-2 encoding, H.264 encoding, H.265 encoding, VC-1 encoding, and synthetic encodings such as Scalable Vector Graphics (SVG) and LASER (ISO/IEC 14496-20), and so forth. It is appreciated that an encoded data stream generally requires more processing and typically more time to read than a data stream which is not encoded. Any recipient of encoded data, whether or not the recipient of the encoded data is the intended recipient, is, at least in potential, able to read encoded data without requiring cryptanalysis. It is appreciated that encoding may be performed in several stages and may include a number of different processes, including, but not necessarily limited to: compressing the data; transforming the

data into other forms; and making the data more robust (for instance replicating the data or using error correction mechanisms).

[0016] The term “compressed” is used throughout the present specification and claims, in all of its grammatical forms, to refer to any type of data stream compression. Compression is typically a part of encoding and may include image compression and motion compensation. Typically, compression of data reduces the number of bits comprising the data. In that compression is a subset of encoding, the terms “encoded” and “compressed”, in all of their grammatical forms, are often used interchangeably throughout the present specification and claims.

[0017] Similarly, the terms “decoded” and “decompressed” are used throughout the present specification and claims, in all their grammatical forms, to refer to the reverse of “encoded” and “compressed” in all their grammatical forms.

[0018] The terms “scrambled” and “encrypted”, in all of their grammatical forms, are used interchangeably throughout the present specification and claims to refer to any appropriate scrambling and / or encryption methods for scrambling and / or encrypting a data stream, and / or any other appropriate method for intending to make a data stream unintelligible except to an intended recipient(s) thereof. Well known types of scrambling or encrypting include, but are not limited to DES, 3DES, and AES. Similarly, the terms “descrambled” and “decrypted” are used throughout the present specification and claims, in all their grammatical forms, to refer to the reverse of “scrambled” and “encrypted” in all their grammatical forms.

[0019] Pursuant to the above definitions, the terms “encoded”; “compressed”; and the terms “scrambled” and “encrypted” are used to refer to different and exclusive types of processing. Thus, a particular data stream may be, for example:

[0020] encoded, but neither scrambled nor encrypted;

[0021] compressed, but neither scrambled nor encrypted;

[0022] scrambled or encrypted, but not encoded;

[0023] scrambled or encrypted, but not compressed;

[0024] encoded, and scrambled or encrypted; or

[0025] compressed, and scrambled or encrypted.

5 [0026] Likewise, the terms “decoded” and “decompressed” on the one hand, and the terms “descrambled” and “decrypted” on the other hand, are used to refer to different and exclusive types of processing.

[0027] By way of introduction, peak signal-to-noise ratio (PSNR) is a widely used quality metric for images and video. Despite the many limitations
10 it has (for example, PSNR works poorly across different video content), it remains an effective tool for quality evaluation for lossy encoding.

[0028] Peak signal-to-noise ratio (PSNR), is the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation. It should be noted that the maximum possible
15 power is not the power of the original signal, but the maximum power over all possible signals. For example, an 8-bit picture pixel can have a maximum value of 255, so the maximum possible power of the 8-bit picture pixel is 255^2 .

[0029] Another category of metrics, linked to encoding, attempts to deal with the distortion incurred by lossy encoding. This second category of
20 metrics typically uses the quantization parameter (QP) to indicate the quality. However, since human vision is less sensitive to distortion in high-complexity/high motion scenes, the QP values are generally normalized by complexity and motion of the video content. Different metrics in this category may differ in how the complexity and motion components are calculated, but the
25 general principles remain basically the same. These metrics are typically easier to compute and generally do not incur too much additional computation than the encoding itself. An example of another metric in this category includes the constant rate factor (CRF) used in x264 (x264 is a general public license free software library for encoding video streams into H.264/MPEG-4). This class of

metrics does not deal with distortion due to resampling so that two videos with the same content but very different spatial resolution may have the same score.

[0030] The complexity/motion-normalized QP is a reasonably good measure for modeling the subjective quality due to lossy encoding and works well for comparing video quality across different content. However, as stated above, complexity/motion-normalized QP is unable to take into account the distortion due to resampling, as videos with different resolutions may end up with different subjective quality after up-sampling, even if they have the same complexity/motion and are encoded with the same QP.

[0031] Reference is now made to Fig. 1, which is a block diagram view of a video processing system 10 constructed and operative in accordance with an embodiment of the present invention.

[0032] The video processing system 10 includes a processor 12 and a memory 14 to store data used by the processor 12.

[0033] The video processing system 10 is operative to calculate a resolution normalized video quality metric based on a combination of a peak-signal-to-noise ratio (PSNR) and a video metric M (such as QP) or a complexity-motion normalized video metric M, for example, but not limited to, a complexity-motion normalized quality metric based on QP (such as constant rate factor (CRF).

[0034] The calculation of the resolution normalized video quality metric is described in more detail with reference to Figs. 2-5, but in overview, the processor 12 of the video processing system 10 calculates the resolution normalized video quality metric of a video portion (either a single video frame or a sequence of video frames) based on an up-sampled PSNR (uPSNR) of the video portion and a mapping between PSNR and the (complexity-motion normalized) video metric M at full resolution for the same video portion.

[0035] In some implementations, the transcoding device may actually encode a lightly compressed “mezzanine” format (somewhere between the original video and the encoded video encoded for serving/broadcasting to the clients) rather than the actual, full-reference original. So for example, a live program might be taken from the cameras at almost 1Gbps and then compressed

by lossy compression to get the bitrate down to, say, 30Mbps before the video is sent to the transcoders. This lightly compressed format is still of a much higher quality than the quality that will be sent to the end-user, but it does not have the same quality as the full-reference original. Nevertheless, it is possible to use this
5 lightly compressed input to the encoders as if it were the actual original full-reference in order to estimate the quality degradation that is occurring during transcoding. Therefore, the video portion may be the original full reference video or a compressed version of the video.

[0036] Up-sampled PSNR (uPSNR), is defined as the PSNR
10 where the corrupting noise is measured by comparing the original signal (the video portion at full resolution (both spatial resolution and frame rate)) with an up-sampled decoded video portion (being the video portion encoded and then decoded and then up-sampled to that full resolution). The up-sampling process typically involves the selection of an interpolation function, for example, but not limited to,
15 bicubic, lanczos, spline, which may result in some marginal differences in the reconstructed picture.

[0037] The inventors have found that uPSNR is a good predictor of subjective quality when comparing different lossy encoding/resampling configurations for the same content. However, when comparing across different
20 content (e.g. a high-motion scene vs. a static scene), uPSNR poorly correlates with subjective quality.

[0038] The inventors have also found that for a given video content, since uPSNR is a good predictor of the subjective quality, it may be used as a reference to normalize a video quality metric such as QP or a
25 complexity/motion-normalized video quality metric based on QP such as CRF across different resolutions. For example, suppose we have two videos of different encoding/resampling configurations, one at the same full resolution as the original video, and the other at a lower resolution. If both videos have the same uPSNR, they may be regarded as having equal subjective quality. The (complexity/motion-
30 normalized video) quality metric value of the full-resolution video may then be

used as the quality score for the lower-resolution video as well in order to provide a resolution normalized video quality metric.

[0039] The resolution normalized M value may be used with live HTTP-based adaptive streaming (HAS) or any other suitable streaming or content provisioning application. In HAS, a video stream is typically divided into segments of a few seconds each. Each segment of the source video is pre-encoded into multiple bitrate representations and stored at a server. Each client adaptively fetches the bitrate segments according to the bandwidth available for downloading content. The different bitrate representations may be generated from the original video via lossy encoding (essentially, by applying a coarser quantizer to the transform coefficients), resampling (including both lowering spatial resolution and lowering frame rate, followed by up-sampling before displaying on screen), or by any suitable combination of the above.

[0040] Another aspect of HAS is that, since video is typically delivered over lossless TCP, the only types of distortion generally expected in the video delivered are those caused by lossy encoding and resampling. Thus, the video quality score can be pre-computed at the same time as encoding the video portion.

[0041] The resolution normalized M value, may be determined for each bitrate segment. The resolution normalized M value is a full-reference metric and utilizes the original lossless video or any relatively high quality lightly encoded video in lieu of the original lossless video (available at the encoder) and is typically robust against varying video content, lossy encoding and resampling.

[0042] Calculation of the resolution normalized video metric is now described in more detail with reference to Figs. 2-5.

[0043] Reference is now made to Fig. 2, which is a chart showing mappings 22, 24 between uPSNR (axis 18) and M (axis 20) at various resolutions for use in the system 10 of Fig. 1.

[0044] The mappings shown in Fig. 2 are for a particular video portion where M is complexity-motion normalized QP.

[0045] Fig. 2 shows a mapping 22 between uPSNR and M at full resolution, Rmax.

[0046] Fig. 2 also shows other mappings 24 between uPSNR and M at resolutions less than full resolution, for example, a mapping 26 between uPSNR and M at resolution Rx (Rx being less than Rmax) is shown. The other mappings 24 are shown for the sake of completeness. However, it will be seen that the other mappings 24 do not need to be determined in order to determine the resolution normalized M value.

[0047] The mapping 22 approximates a straight line mapping and therefore, the mapping between uPSNR and M at the full resolution (Rmax) may be determined based on determining two uPSNR values (PSNR₁ and PSNR₂) and corresponding values for M (V₁ and V₂). It will be appreciated that the mapping between uPSNR and M at full resolution is the same as the mapping between PSNR and M at full resolution because the video portion does not need to be up-sampled since it is already at full resolution.

[0048] Where the mapping between uPSNR and M at full resolution, Rmax, is expected to be a non-linear relationship then the mapping between uPSNR and M at full resolution, Rmax, may be determined based on determining a suitable number of values of uPSNR and corresponding M values and then applying a suitable curve fitting algorithm in order to determine the mapping between uPSNR and M at full resolution.

[0049] Reference is now made to Fig. 3, which is a flow chart type view showing a method of determining the mapping 22 (Fig. 2) between PSNR and M at full resolution for use in the system 10 of Fig. 1.

[0050] The processing performed by processor 12 (Fig. 1) of the video processing system 10 (Fig. 1) may be divided into sub-processors including a video encoder 34, a video decoder 28, a PSNR calculation unit 30, a PSNR vs. M mapping unit 32. Some or all of the sub-processors may be combined into a single processing unit or be implemented as separate processing units operationally connected to each other.

[0051] The video encoder 34 is operative to encode a video portion 36 at a first video quality and at full resolution R_{max} yielding an encoded video portion 38.

[0052] The video decoder 28 is operative to decode the encoded video portion 38 yielding a decoded video portion 40.

[0053] The PSNR calculation unit 30 is operative to receive the decoded video portion 40 and the video portion 36 and calculate a first peak signal to noise ratio, $PSNR_1$ (block 42) for the encoded version of the video portion 36 encoded at the first video quality and R_{max} . The calculation of $PSNR_1$ includes comparing the decoded video portion 40 (being a decoded version of the encoded version of the video portion 36) with the (original or lightly compressed) video portion 36 to determine the corrupting noise. It should be noted that each PSNR value calculation described herein includes comparing a decoded version of the encoded version of the video portion 36 to the (original or lightly compressed) video portion 36. Similarly, each uPSNR value calculation described herein includes comparing an up-sampled (to R_{max}) decoded version of the encoded version of the video portion 36 to the video portion 36.

[0054] The processor 12 is operative to determine a first M value, V_1 , for the encoded version of the video portion 36 encoded at the first video quality and R_{max} based on an encoding parameter such as QP and optionally on encoding parameters such as complexity and motion of the encoded video portion 38 (block 44).

[0055] The above steps are repeated wherein, the video portion 36 is also encoded by the video encoder 34 at a second video quality, at full resolution R_{max} , and then decoded by the video decoder 28 and then processed by the PSNR calculation unit 30 to calculate a second peak signal to noise ratio, $PSNR_2$ (block 42) for the encoded version of the video portion 36 encoded at the second video quality and R_{max} .

[0056] The processor 12 is also operative to calculate a second M value, V_2 , for the encoded version of the video portion 36 encoded at the second video quality and R_{max} (block 44).

[0057] The above processing supplies the values $PSNR_1$ and
5 $PSNR_2$ and the corresponding values V_1 and V_2 .

[0058] The PSNR vs. M mapping unit 32 is operative to determine the mapping (block 46) between the peak signal to noise ratio, PSNR, and the video quality metric, M, for the encoded version of the video portion 36 when the encoded version is encoded at full resolution R_{max} based on V_1 , V_2 ,
10 $PSNR_1$ and $PSNR_2$.

[0059] The example of Fig. 3 assumes that there is a straight line mapping between PSNR and M at full resolution R_{max} . If the relationship is non-linear, more than two points may be needed to provide a sufficient mapping. If the relationship is non-linear, a linear relationship may be assumed in some cases for
15 simplicity purposes but may distort the calculated metrics.

[0060] It should be noted that the mapping between PSNR and M is determined for a specific video portion and is repeated for each video portion that a quality metric needs to be calculated for.

[0061] Reference is now made to Fig. 4, which is a view of the
20 chart of Fig. 2 showing looking up $uPSNR(R_x)$ value in the mapping to determine a resolution normalized M value.

[0062] Fig. 4 shows the mapping 22 between $uPSNR$ and M at full resolution, R_{max} , for a particular video portion.

[0063] The particular video portion is encoded at a resolution R_x
25 and has a $uPSNR$ value $uPSNR(R_x)$ and a (complexity-motion normalized) metric M value of V.

[0064] The (complexity-motion normalized) metric M value V is resolution normalized by looking up the value $uPSNR(R_x)$ in the mapping 22

between uPSNR and M at full resolution Rmax, giving the value V' which is the resolution normalized value of V.

[0065] The calculation of V' is now described in more detail with reference to Fig. 5.

5 [0066] Reference is now made to Fig. 5, which is a flow chart type view showing a method to determine the resolution normalized M value, V' of Fig. 4.

[0067] As described above with reference to Fig. 3, the processing performed by processor 12 (Fig. 1) of the video processing system 10 (Fig. 1) may be divided into sub-processors. In addition to the sub-processors listed in Fig. 3, the sub-processors may include an interpolation unit 48 and a look-up unit 50. Some or all of the sub-processors may be combined into a single processing unit or be implemented as separate processing units operationally connected to each other.

15 [0068] The processor 12 (Fig. 1) is operative to determine an up-sampled peak signal to noise ratio, uPSNR(Rx), of an up-sampled decoded video portion 52 based on the steps below.

[0069] The video decoder 28 is operative to receive an encoded video portion 54 (encoded at resolution Rx) and decode the encoded video portion 54 yielding a decoded video portion 56. The interpolation unit 48 is operative to up-sample the decoded video portion 56 to the full resolution Rmax yielding the up-sampled decoded video portion 52. Therefore, the up-sampled decoded video portion 52 derives from the video portion 36 being encoded at a resolution Rx, which is less than the full resolution Rmax, yielding the encoded video portion 54 and the encoded video portion 54 being decoded and then up-sampled to the full resolution Rmax.

[0070] Up-sampling typically includes changing a higher spatial resolution and/or changing to a higher frame rate. Changing to a higher frame rate could be implemented by generating intermediate frame(s) between existing frames based on temporal interpolation, and compared with an original video portion.

[0071] The PSNR calculation unit 30 is operative to receive, by being read from the memory 14 (Fig. 1) the video portion 36 and the up-sampled decoded video portion 52. The PSNR calculation unit 30 is operative to calculate the uPSNR(Rx) value based on comparing the inputs including the video portion 36 and the up-sampled decoded video portion 52 (block 58).

[0072] The processor 12 (Fig. 1) is operative to read, from the memory 14 (Fig. 1), the mapping 22 between the peak signal to noise ratio, PSNR, and the video quality metric, M, for the encoded version of the video portion 36 encoded at full resolution Rmax as depicted in Figs. 2 and 4.

[0073] The look-up unit 50 is operative to read the value uPSNR(Rx), from the memory 14 (Fig. 1), and look-up a value, V' (block 60), mapping to uPSNR(Rx) in the mapping 22 between PSNR and M at full resolution, Rmax. V' represents a resolution normalized M value of the encoded video portion 54 encoded at the resolution Rx.

[0074] The video portion 36 and encoded video portion 54 may be a single video frame comprised in a video sequence (for example, a bitrate segment in an adaptive bitrate (ABR) system). The video sequence includes a plurality of video frames (not shown). The processor 12 (Fig. 1) is optionally operative to determine a resolution normalized M value for an encoded version of each of the plurality of video frames (each encoded at the same resolution Rx) using the method described above with reference to Figs. 2-5 for each of the video frames (block 62). It should be noted that a new mapping 22 may be determined for each of the video frames using the method described above with reference to Fig. 3. The processor 12 (Fig. 1) is operative to calculate a resolution normalized M value for an encoded version of the video sequence based on the resolution normalized M value of each of the plurality of video frames (block 64), for example by determining an average of all the resolution normalized M values for the video frames of the video sequence or take a percentile score (e.g. 10%-worst score) over all the video frames within the video sequence.

[0075] Reference is now made to Fig. 6, which is a block diagram view of a client device 66 constructed and operative in accordance with an embodiment of the present invention.

[0076] The client device 66 includes a processor 68 and a
5 memory 70 to store data used by the processor 68.

[0077] Reference is now made to Fig. 7, which is a flow chart showing a method of selecting bitrate segments in the client device 66 of Fig. 6.

[0078] The processor 68 (Fig. 6) is operative to receive a choice of a plurality of different bitrate segments for download (block 72).

10 [0079] The processor 68 (Fig. 6) is operative to receive a resolution normalized video quality metric V' for each of the different bitrate segments (block 74) from the video processing system 10 or any other system which prepared and/or stored the resolution normalized video quality metric. The resolution normalized video quality metric V' for each of the different bitrate
15 segments is determined by the method described above with reference to Figs. 1-5.

[0080] The processor 68 (Fig. 6) is operative to select one of the different bitrate segments for download based on the resolution normalized video quality metric V' for each of the different bitrate segments (block 76).

20 [0081] The processor 68 (Fig. 6) is operative to download the selected one of the different bitrate segments (block 78).

[0082] Reference is now made to Fig. 8, which is a block diagram view of a device 80 constructed and operative in accordance with an embodiment of the present invention.

25 [0083] The device 80 includes a processor 82 and a memory 84 to store data used by the processor 82. The device 80 is operative to calculate video quality metrics for the video rendered by client devices including the client device 66. The device 80 may be implemented in a content server or in another device

[0084] Reference is now made to Fig. 9, which is a flow chart showing a method of calculating a video quality metric by the device 80 of Fig. 8 for the segments selected by the client device 66 of Fig. 6.

[0085] The processor 82 of the device 80 (Fig. 8) is operative to
5 receive data of which different bitrate segments of content were selected for download by the client device 66 of Fig. 6 (block 86).

[0086] The processor 82 of the device 80 (Fig. 8) is operative to receive a resolution normalized video quality metric V' for each of the selected bitrate segments (block 88). The resolution normalized video quality metric V' for
10 each of the selected bitrate segments being determined by the method described above with reference to Figs. 1-5.

[0087] The processor 82 of the device 80 (Fig. 8) is operative to calculate an overall quality metric (for example, but not limited to, arithmetic mean or percentile) based on the data of which different bitrate segments of the
15 content were selected for download by the client device and the resolution normalized video quality metric for each of the selected bitrate segments (block 90).

[0088] In practice, some or all of the functions described hereinabove may be combined in a single physical component or, alternatively,
20 implemented using multiple physical components. These physical components may comprise hard-wired or programmable devices, or a combination of the two. In some embodiments, at least some of the functions of the processing circuitry may be carried out by a programmable processor under the control of suitable software. This software may be downloaded to a device in electronic form, over a
25 network, for example. Alternatively or additionally, the software may be stored in tangible, non-transitory computer-readable storage media, such as optical, magnetic, or electronic memory.

[0089] It is appreciated that software components may, if desired, be implemented in ROM (read only memory) form. The software components
30 may, generally, be implemented in hardware, if desired, using conventional techniques. It is further appreciated that the software components may be

instantiated, for example: as a computer program product or on a tangible medium. In some cases, it may be possible to instantiate the software components as a signal interpretable by an appropriate computer, although such an instantiation may be excluded in certain embodiments of the present invention.

5 [0090] It will be appreciated that various features of the invention which are, for clarity, described in the contexts of separate embodiments may also be provided in combination in a single embodiment. Conversely, various features of the invention which are, for brevity, described in the context of a single embodiment may also be provided separately or in any suitable sub-combination.

10 [0091] It will be appreciated by persons skilled in the art that the present invention is not limited by what has been particularly shown and described hereinabove. Rather the scope of the invention is defined by the appended claims and equivalents thereof.

What is claimed is:

CLAIMS

1. A video processing system comprising: a processor; and a memory to store data used by the processor, wherein the processor is operative to:
 - 5 read, from the memory, a mapping between a peak signal to noise ratio, PSNR, and a video quality metric, M, for an encoded version of a first video portion encoded at a full resolution Rmax;
 - read, from the memory, an up-sampled peak signal to noise ratio, uPSNR(Rx), of an up-sampled decoded video portion, the up-sampled decoded
 - 10 video portion deriving from the first video portion being encoded at a resolution Rx, which is less than the full resolution Rmax, yielding an encoded video portion, the encoded video portion being decoded and then up-sampled to the full resolution Rmax; and
 - lookup a value, V', that maps to uPSNR(Rx) in the mapping, V' representing a resolution normalized M value of the encoded video portion
 - 15 encoded at the resolution Rx.
2. The system according to claim 1, wherein the first video portion is a video frame comprised in a video sequence including a plurality of video frames, wherein the processor is operative to:
 - 20 determine a resolution normalized M value for an encoded version of each of the plurality of video frames; and
 - calculate a resolution normalized M value for an encoded version of the video sequence based on the resolution normalized M value of each of the plurality of video frames.
- 25 3. The system according to claim 1 or claim 2, wherein the processor is operative to:
 - receive the first video portion;

read, from the memory, the up-sampled decoded video portion; and
calculate $uPSNR(R_x)$ based on comparing inputs including the first
video portion and the up-sampled decoded video portion.

4. The system according to any of claims 1-3, wherein the processor is
5 operative to:
- receive the encoded video portion encoded at resolution R_x ;
decode the encoded video portion yielding a decoded video portion;
and
up-sample the decoded video portion to the full resolution R_{max}
10 yielding the up-sampled decoded video portion.

5. The system according to any of claims 1-4, wherein the processor is
operative to determine the mapping between the peak signal to noise ratio, PSNR,
and the video quality metric, M , for the encoded version of the first video portion
when the encoded version is encoded at full resolution R_{max} .

- 15 6. The system according to claim 5, wherein the processor is operative
to:
- encode the first video portion at a first video quality, at full
resolution R_{max} ;
calculate a first peak signal to noise ratio, $PSNR_1$, for the encoded
20 version of the first video portion encoded at the first video quality and R_{max} ;
calculate a first M value, V_1 , for the encoded version of the first
video portion encoded at the first video quality and R_{max} ;
encode the first video portion at a second video quality, at full
resolution R_{max} ;
25 calculate a second peak signal to noise ratio, $PSNR_2$, for the
encoded version of the first video portion encoded at the second video quality and
 R_{max} ; and
calculate a second M value, V_2 , for the encoded version of the first
video portion encoded at the second video quality and R_{max} ;

determine the mapping between the peak signal to noise ratio, PSNR, and the video quality metric, M, for the encoded version of the first video portion when the encoded version is encoded at full resolution R_{max} based on V1, V2, PSNR1 and PSNR2 as input.

5 7. The system according to any of claims 1-6, wherein the M is a complexity-motion normalized quality metric or a quantization parameter.

8. A video processing method comprising:

reading, from a memory, a mapping between a peak signal to noise ratio, PSNR, and a video quality metric, M, for an encoded version of a first video
10 portion encoded at full resolution R_{max} ;

reading, from a memory, an up-sampled peak signal to noise ratio, $uPSNR(R_x)$, of an up-sampled decoded video portion, the up-sampled decoded video portion deriving from the first video portion being encoded at a resolution R_x , which is less than the full resolution R_{max} , yielding an encoded video portion,
15 the encoded video portion being decoded and then up-sampled to the full resolution R_{max} ; and

looking-up a value, V' , that maps to $uPSNR(R_x)$ in the mapping, V' representing a resolution normalized M value of the encoded video portion encoded at the resolution R_x .

20 9. The method according to claim 8, wherein the first video portion is a video frame comprised in a video sequence including a plurality of video frames, wherein the method further comprises:

determining a resolution normalized M value for an encoded version of each of the plurality of video frames; and

25 calculating a resolution normalized M value for an encoded version of the video sequence based on the resolution normalized M value of each of the plurality of video frames.

10. The method according to claim 8 or claim 9, wherein the M is a complexity-motion normalized quality metric or a quantization parameter.
11. A client device, comprising: a processor; and a memory to store data used by the processor, wherein the processor is operative to:
- 5 receive a choice of a plurality of different bitrate segments for download;
- receive a resolution normalized video quality metric for each of the different bitrate segments, the resolution normalized video quality metric for each of the different bitrate segments being determined by the steps of the method of
- 10 claim 8;
- select one of the different bitrate segments for download based on the resolution normalized video quality metric for each of the different bitrate segments; and
- download the selected one of the different bitrate segments.
- 15 12. The client device according to claim 11, wherein the M is a complexity-motion normalized quality metric or a quantization parameter.
13. A device, comprising: a processor; and a memory to store data used by the processor, wherein the processor is operative to:
- receive data of which different bitrate segments of content were
- 20 selected for download by a client device;
- receive a resolution normalized video quality metric for each of the selected bitrate segments, the resolution normalized video quality metric for each of the selected bitrate segments being determined by the steps of the method of claim 8; and
- 25 calculate an overall quality metric based on the data of which different bitrate segments of the content were selected for download by the client device and the resolution normalized video quality metric for each of the selected bitrate segments.

14. The device according to claim 13, wherein the M is a complexity-motion normalized quality metric or a quantization parameter.
15. A method comprising:
receiving a choice of a plurality of different bitrate segments for
5 download;
receiving a resolution normalized video quality metric for each of
the different bitrate segments, the resolution normalized video quality metric for
each of the different bitrate segments being determined by the steps of the method
of claim 8;
10 selecting one of the different bitrate segments for download based
on the resolution normalized video quality metric for each of the different bitrate
segments; and
downloading the selected one of the different bitrate segments.
16. The method according to claim 15, wherein the M is a complexity-
15 motion normalized quality metric or a quantization parameter.
17. A method comprising:
receiving data of which different bitrate segments of content were
selected for download by a client device;
receiving a resolution normalized video quality metric for each of
20 the selected bitrate segments, the resolution normalized video quality metric for
each of the selected bitrate segments being determined by the steps of the method
of claim 8; and
calculating an overall quality metric based on the data of which
different bitrate segments of the content were selected for download by the client
25 device and the resolution normalized video quality metric for each of the selected
bitrate segments.
18. The device according to claim 17, wherein the M is a complexity-
motion normalized quality metric or a quantization parameter.

1/7

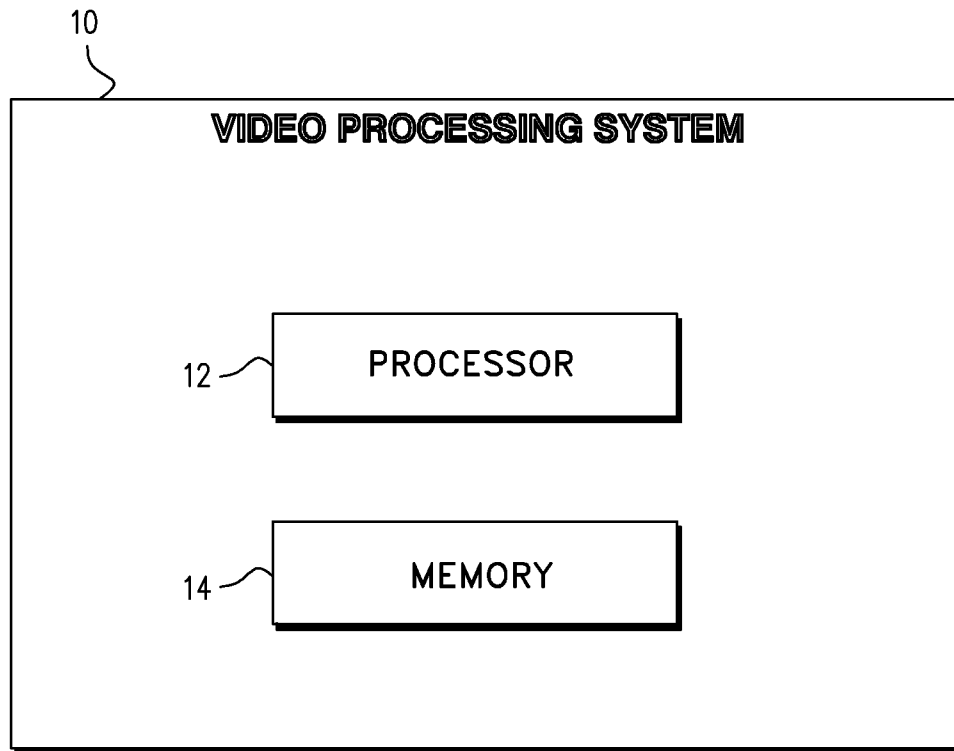
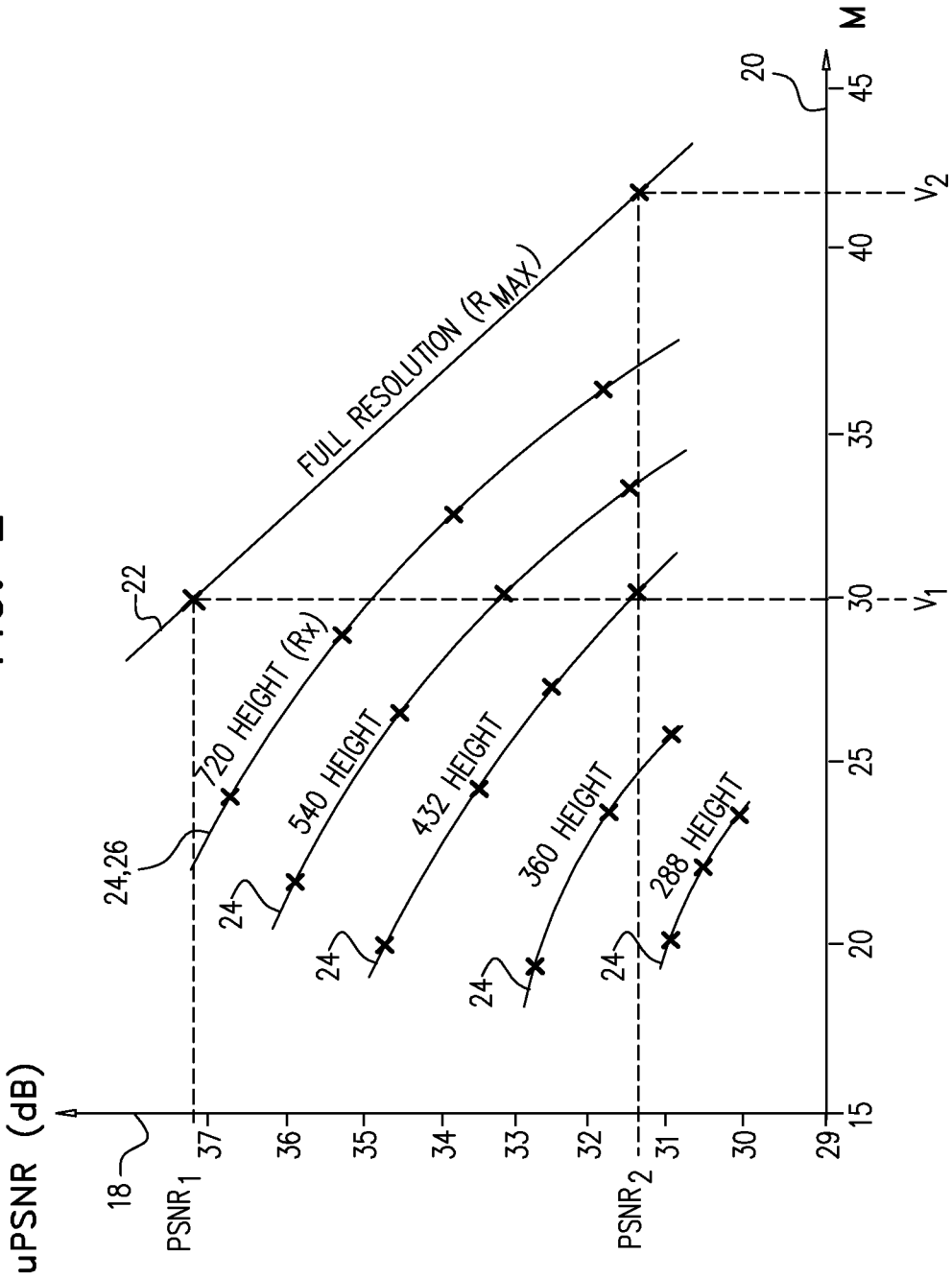


FIG. 1

FIG. 2



3/7

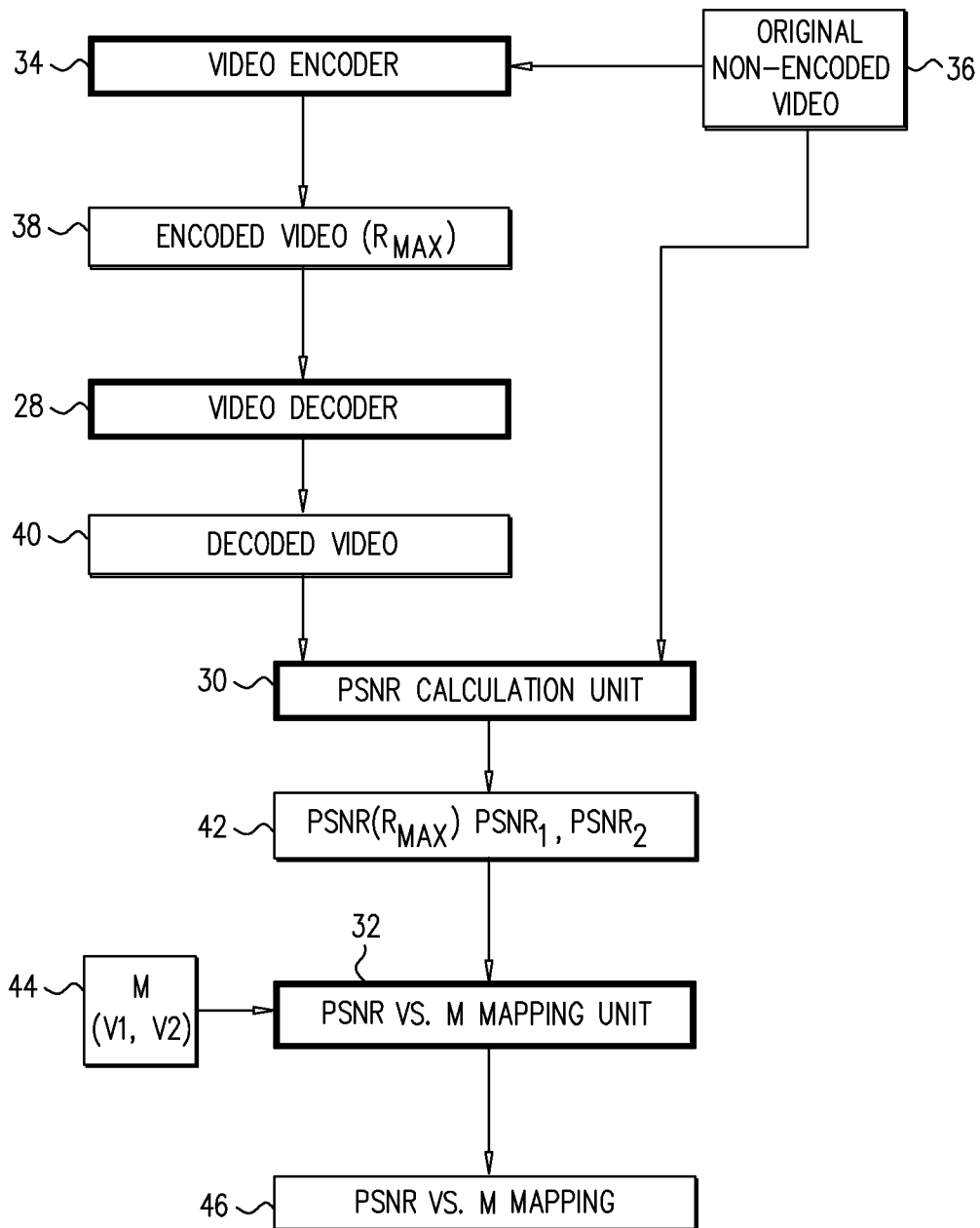
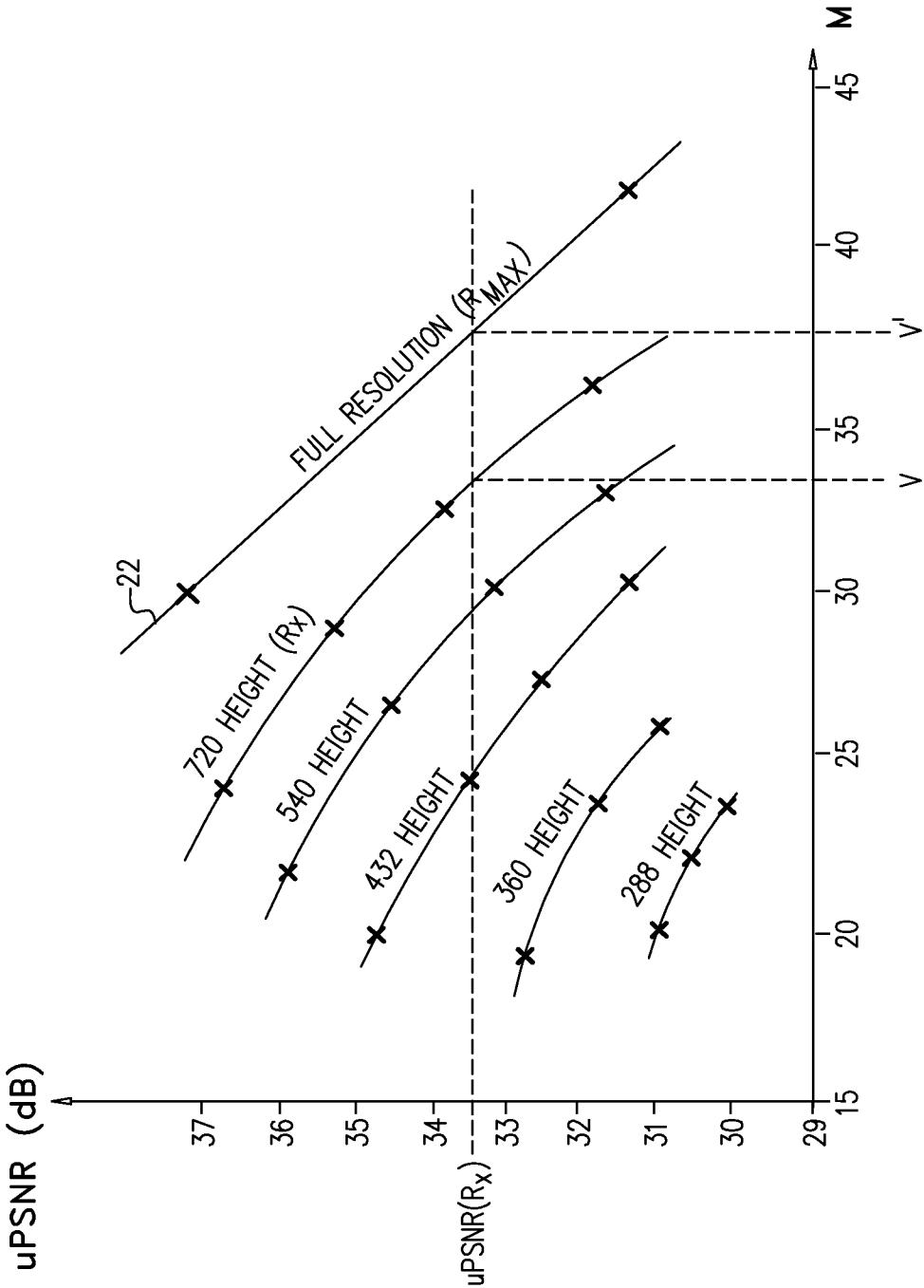


FIG. 3

FIG. 4



5/7

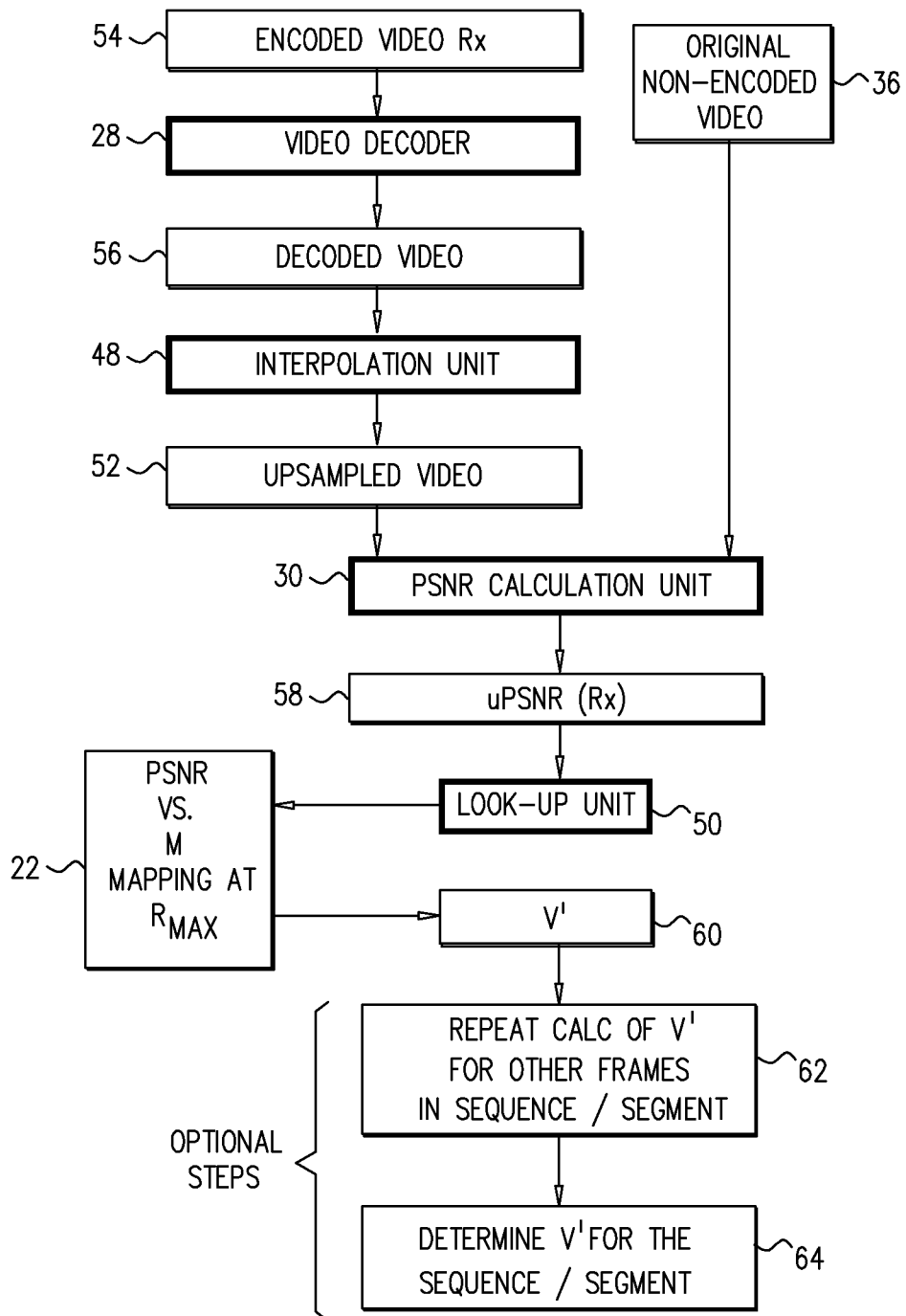


FIG. 5

6/7

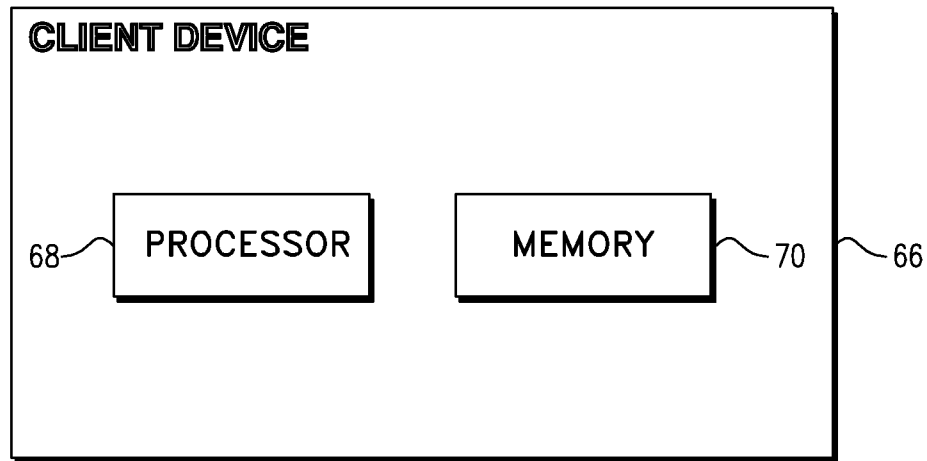


FIG. 6

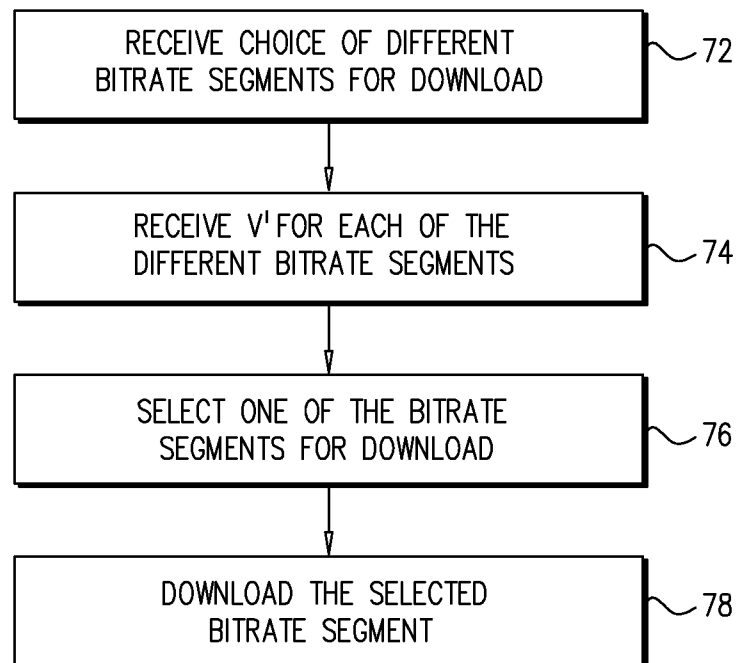


FIG. 7

7/7

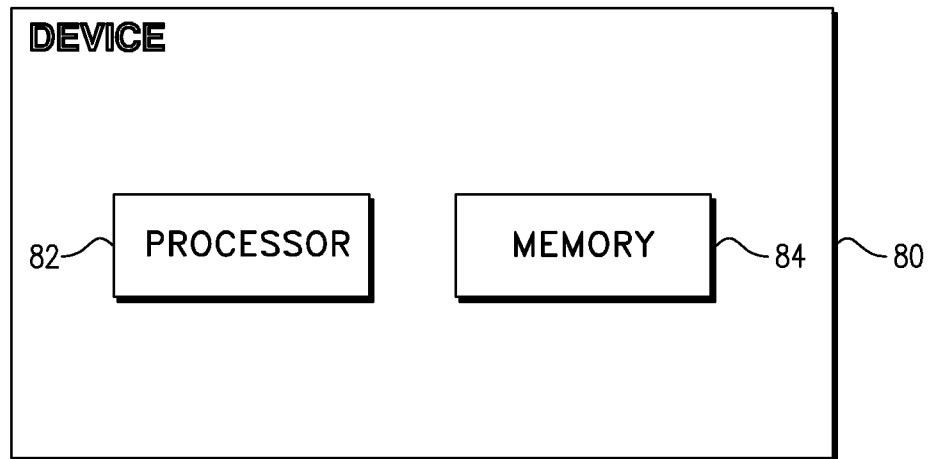


FIG. 8

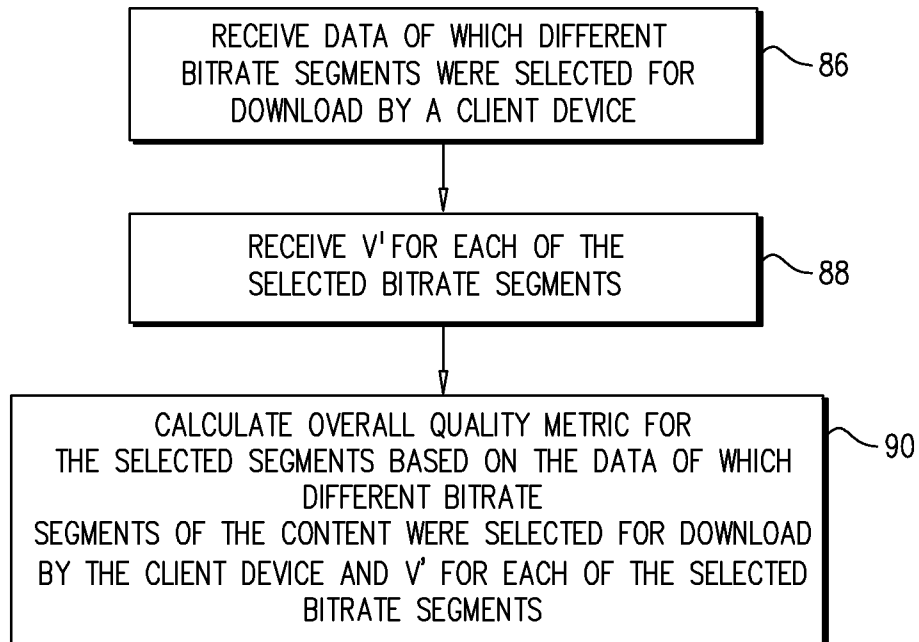


FIG. 9

INTERNATIONAL SEARCH REPORT

International application No
PCT/IB2015/054313

A. CLASSIFICATION OF SUBJECT MATTER
INV. H04N17/00 H04N19/154
ADD.

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

B. FIELDS SEARCHED

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

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

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2014/201382 A1 (SHIVADAS ABHISHEK [US] ET AL) 17 July 2014 (2014-07-17)	1,2,4,5,8-10
A	paragraph [0050] - paragraph [0059]	3,6,7
Y	----- RIES M ET AL: "Subjective evaluation of video quality for H.264 encoded sequences", MOBILE FUTURE, 2004 AND THE SYMPOSIUM ON TRENDS IN COMMUNICATIONS. SYM POTIC '04. JOINT 1ST WORKSHOP ON BRATISLAVA, SLOVAKIA 24-26 OCT. 2004, PISCATAWAY, CA, USA, IEEE, US, 24 October 2004 (2004-10-24), pages 191-194, XP032371378, DOI: 10.1109/TIC.2004.1409531 ISBN: 978-0-7803-8556-6 page 192 - column 2 -----	1,2,4,5,8-10



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

27 July 2015

Date of mailing of the international search report

07/08/2015

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Penchev, Petyo

INTERNATIONAL SEARCH REPORT

International application No.
PCT/IB2015/054313

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☒ Claims Nos.: 11-18
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
see FURTHER INFORMATION sheet PCT/ISA/210
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box II.2

Claims Nos.: 11-18

Independent claims 11, 13 and 15, 17 refer to use of a resolution normalized quality metric generated in claim 8. Such a reference renders the scope of the claims unclear, because it is impossible to determine whether the resolution normalized quality metric is determined by the method claim 8 or some other method.

The applicant's attention is drawn to the fact that claims relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure. If the application proceeds into the regional phase before the EPO, the applicant is reminded that a search may be carried out during examination before the EPO (see EPO Guidelines C-IV, 7.2), should the problems which led to the Article 17(2) declaration be overcome.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/IB2015/054313

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2014201382	A1	17-07-2014	NONE
