

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
2 February 2012 (02.02.2012)

PCT

(10) International Publication Number
WO 2012/013777 A2

(51) International Patent Classification:
H04N 7/26 (2006.01) *H04N 17/00* (2006.01)
H04N 7/50 (2006.01)

(74) Agent: **VOSSIUS & PARTNER**; Siebertstraße 4, 81675 München (DE).

(21) International Application Number:
PCT/EP2011/063091

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

(22) International Filing Date:
29 July 2011 (29.07.2011)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
10171448.3 30 July 2010 (30.07.2010) EP

(71) Applicants (*for all designated States except US*):
DEUTSCHE TELEKOM AG [—/DE]; Friedrich-Ebert-Allee 140, 53113 Bonn (DE). **TECHNISCHE UNIVERSITÄT BERLIN** [—/DE]; Strasse des 17. Juni 135, 10623 Berlin (DE).

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, 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, ML, MR, NE, SN, TD, TG).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **ARGYROPOULOS, Savvas** [GR/DE]; Krumme Straße 74, 10585 Berlin (DE). **FEITEN, Bernhard** [DE/DE]; Am Triftpark 13, 13437 Berlin (DE). **GARCIA, Marie-Neige** [FR/DE]; Pintschstraße 4, 10249 Berlin (DE). **LIST, Peter** [DE/DE]; Beethovenstraße 15, 64859 Eppertshausen (DE). **RAAKE, Alexander** [DE/DE]; Chodowieckstraße 23, 10405 Berlin (DE).

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*



WO 2012/013777 A2

(54) Title: METHOD AND APPARATUS FOR ASSESSING THE QUALITY OF A VIDEO SIGNAL DURING ENCODING OR COMPRESSING OF THE VIDEO SIGNAL

(57) Abstract: The invention provides a method and apparatus for assessing the quality of a video signal during encoding or compressing of the video signal, the method comprising the steps of: a) estimating the quality of the encoded or compressed video signal using one or more parameters; and b) using the length of the Group of Pictures, GOP, and/or GOP-length-related parameters as additional parameter(s) to adjust the estimated video signal quality.

- Reduced-Reference (RR): partial information extracted from the source signal is required.
- No-Reference (NR): no reference signal is required.

5 *Input parameters types*

- signal/media-based: the decoded image (pixel-information) is required.
- parameter-based: bitstream-level information is required. Information can range from packet-header information, requiring parsing of the packet-headers but not (full- or partial) decoding of the bitstream to the complete decoding of the bitstream.

10

Type of application

- Network Planning: the model or measurement system is used before the implementation of the planning in order to plan the best possible implementation.
- Service Monitoring: the model is used during service operation.

15

Related information of the types of video quality models can be found in references [1], [2], or [3].

20

In the context of MPEG-based video services, one of the parameters influencing the video perceived quality is the GOP-Structure (GOP = Group of Pictures), including the GOP-length, i.e., the distance between frames which do not require previous or further frames to be decoded, the so-called ‘key-frames’ or ‘I-frames’. One Group of Picture covers one I-frame and all frames till the next I-frame of the video sequence.

25

The GOP-structure – and thus GOP-length - is generally chosen as a trade-off between encoding efficiency and error-propagation (see, e.g., references [4], [5], [6]). In these references, the authors provide guidelines for selecting the most appropriate GOP structure for MPEG.

30

Some models take as input parameters GOP-related parameters but only under packet loss conditions, as in references [2], [7], [8], [9], or [10]. However, they consider only fixed GOP lengths and examine the impact on quality based on the temporal distance of the frame where the packet loss occurs to the next key frame. The quality impact of the GOP-length on encoding is not taken into account.

Quality estimation methods commonly support a distinguished estimation of the quality related to the coding (compression, Q_{cod}) of the video signal and the quality due to packet loss during transmission (Q_{trans}). Quality estimation methods commonly use one of two approaches to combine an estimation concerning the quality of the compression and the transmission quality. Equation (1) and (2) illustrate the two different approaches

$$Q = Q0 - Q_{cod} - Q_{trans} \quad , \quad Q0, Qx \quad 0 \dots 100 \quad (1)$$

$$10 \quad Q = Q0 * Q_{cod} * Q_{trans} \quad , \quad Q0, Qx \quad 0 \dots 1 \quad (2),$$

in which $Q0$ represents the base quality or a function of the base quality.

Summary of the Invention

15

According to a first aspect, the invention provides a method for assessing the quality of a video signal during encoding or compressing of the video signal, the method comprising the steps of:

- a) estimating the quality of the encoded or compressed video signal using one or more parameters; and
- 20 b) using the length of the Group of Pictures, GOP, and/or GOP-length-related parameters as additional parameter(s) to adjust the estimated video signal quality.

Thus, the method of the invention focuses on the quality estimation of the term characterizing the compression efficiency Q_{cod} . The method of the invention may be combined with different methods for quality estimation of packet loss in video streams.

The method of the invention is a parameter-based video quality model with light-weight parameters, and thus provides a video quality model suitable both for network planning and service monitoring. In case of network planning, values of the parameters are assumed by the network planner, based on knowledge of previously developed similar networks. In case of service monitoring, the model takes as input parameters extracted from the bitstream. In principle, the measurement system in which the method of the invention is embedded can be placed at different locations in the network. However, the closer the probe to the user device

30

is, the more representative of the actual quality at the user side the predicted quality is, when considering packet loss. In case of service monitoring, the parameters do not require access to the payload, and therefore do not require either a partial- or full-decoding of the bistream. A light-parsing of the packet headers is sufficient for accessing the parameters to be sent to the model, i.e., method. Note that if deeper parsing is allowed, as with un-encrypted data, the parameters can also be used with additional parameters extracted from the bit-stream by partial or full-decoding.

The method of the invention considers the parameters of the encoding process in detail. Known parametric quality estimation methods use the bit-rate, the frame-rate, the video resolution, the codec type and the content type to estimate the quality of a compressed video stream. The invention goes beyond these conventional methods by using the GOP-length or GOP-length-related parameters as an additional parameter for adjusting the estimated video quality. This parameter has a direct impact on the number of bits per frame, and thus on the perceived video quality. As a consequence, the invention considers this parameter as input of the model in addition to the above parameters. With the method of the invention a much more accurate estimation of the related perceived quality can be achieved.

Again, the method of the invention takes as input parameters such as the video resolution, the codec type, the content type, the bit-rate, the frame-rate and the key-frame-rate, and output an estimated video quality (Q_{cod}) based on those parameters. This can be written as in equation (3):

$$Q_{cod} = f(br, fr, cod, cont, res, kfr, I, G) \quad (3)$$

25

in which

br	...	bitrate, the number of bits per second
fr	...	frame-rate, the number of video frames per second
cod	...	employed codec (for instance H.264 main profile)
$cont$...	content type (for instance TVnews, sport, movie, etc.)
res	...	video resolution (for instance High Definition, 1920x1080 pixels progressive)
kfr	...	key-frame-rate, the number of key-frames per second
I	...	number of bits in an I-frame
G	...	number of bits in a GOP

30

Note that for some of the parameters their values might not be used directly into the function, but as a switch for selecting the appropriate coefficients values of the function.

- 5 For instance, a possible model for Q_{cod} in equation (1) for IPTV in which the frame-rate is considered constant is shown in equation (4):

$$Q_{cod} = a * \exp(b * br) + c \quad (4)$$

- 10 In equation (4), the values of the a , b and c coefficients depend on the employed video codec, on the video resolution and on the content type. Those coefficients are preferably obtained in a least-square-error curve fitting procedure using the ratings of perception tests as target values.

- 15 Following the invention, and still using ratings of perception tests, the key-frame-rate is used as additional parameter into equation (4), yielding equation (5):

$$Q_{cod} = (a1 * kfr + a2) * \exp(b * br) + c \quad (5)$$

- 20 The coefficients $a1$, $a2$ and b may be obtained in a least-square-error curve fitting procedure using the ratings of perception tests as target values. $a1$ and $a2$ depend on the given encode settings and content type.

- 25 The key-frame-rate may be obtained by computing the ratio of the frame-rate over the GOP length (see Fig.2).

- In an alternative embodiment of the invention, the coefficients $a1$ and $a2$ can be explicitly dependent on additional information extracted from the bit-stream or packet headers. For example, $a1 = I / G * a1'$ and $a2 = I / G * a2'$, where $a1'$ and $a2'$ are curve-fitting parameters that represent given encoder settings, but excluding a part of the variability which can directly be measured using I / G . Thus, the impact of the GOP-length and GOP-length-related parameters is modulated by the key-frame size in bytes normalized by the number of bytes in the GOP. Following this alternative embodiment, we obtain equation (6):
- 30

$$Q_{cod} = \frac{I}{G} * (a1'kfr * +a2') * \exp(b * br) + c \quad (6)$$

Indeed the influence of the key-frame rate also depends on how the information bits are spread over the GOP. If the information bits are equally spread over the frames of the GOP, and in the considered case of no packet-loss, the key-frame-rate has no influence on the quality. If all information bits are in the I-frame of the GOP, the influence of the key-frame-rate is maximal. This impact of the bit distribution over the GOP is captured by the ratio I/G , where G is the number of bits in a GOP and I is the number of bits in an I-frame.

As explained above, it is preferred according to the invention that the one or more parameter used in step a) of the method is selected from the set comprising: bit rate, frame rate, video resolution, codec type, content type.

The values of the parameters used in step a) may be computed from the packet header information extracted from the bit stream of the video signal and/or derived from side information.

Other aspects, features, and advantages will be apparent from the summary above, as well as from the description that follows, including the figures and the claims.

Brief Description of the Drawings

Fig. 1 shows a block diagram of the general framework of the estimation of quality related to compression; and

Fig. 2 shows a detailed view of the computation of the key-frame-rate.

Description of Detailed Embodiments of the Invention

Fig. 1 shows the general framework for estimating the perceived video quality in case of compression degradation only (Q_{cod}) either in case of network planning or in case of service monitoring.

In the case of network planning (block 100a), values of the parameters (block 200) to be sent to the video quality model (block 300) are estimated by the network planner. In the case of

service monitoring (block 100b), the parameters to be sent to the video quality model are computed from the packet header information extracted from the bit-stream.

Fig. 2 shows a detailed view of the computation of the key-frame-rate. As shown in Fig. 2, one GOP covers one I-frame and all frames till the next I-frame. If n is the number of frames on a time window of t or more seconds, then $fr = n/t$ is the frame rate of the video sequence. If d is the number of frames between two I-frames, then $kfr = fr/d$ is the key-frame-rate of the video sequence.

10 While the invention has been illustrated and described in detail in the drawings and foregoing description, such illustration and description are to be considered illustrative or exemplary and not restrictive. It will be understood that changes and modifications may be made by those of ordinary skill within the scope of the following claims. In particular, the present invention covers further embodiments with any combination of features from different
15 embodiments described above and below.

Furthermore, in the claims the word "comprising" does not exclude other elements or steps, and the indefinite article "a" or "an" does not exclude a plurality. A single unit may fulfil the functions of several features recited in the claims. The terms "essentially", "about",
20 "approximately" and the like in connection with an attribute or a value particularly also define exactly the attribute or exactly the value, respectively. Any reference signs in the claims should not be construed as limiting the scope.

[1] A. Takahashi, D. Hands, and V. Barriac, "Standardization Activities in the ITU for a QoE Assessment of IPTV," in *IEEE Communication Magazine*, 2008.

5 [2] S. Winkler and P. Mohandas, "The Evolution of Video Quality Measurement: From PSNR to Hybrid Metrics," in *IEEE Trans. Broadcasting*, 2008.

[3] A. Raake, M.N. Garcia, S. Moeller, J. Berger, F. Kling, P. List, J. Johann, and C. Heidemann, "T-V-MODEL: Parameter-based prediction of IPTV quality," in *Proc. of ICASSP*, 2008.

10

[4] Huahui Wu, Mark Claypool, and Robert Kinicki, GUIDELINES FOR SELECTING PRACTICAL MPEG GROUP OF PICTURES", In Proceedings of IASTED International Conference on Internet and Multimedia Systems and Applications (EuroIMSA), Innsbruck, Austria, February 2006 <http://web.cs.wpi.edu/~claypool/papers/practical-gop/vGOP.pdf>

15

[5] Árpád Huszák and Sándor Imre, "ANALYSIS GOP STRUCTURE AND PACKET LOSS EFFECTS ON ERROR PROPAGATION IN MPEG-4 VIDEO STREAMS", in Proceedings of the 4th International Symposium on Communications, Control and Signal Processing 2010 (ISCCP 2010), Limassol, Cyprus, 3-5 March 2010

20

[6] Tektronix, "An Analysis of MPEG Encoding Techniques on Picture Quality, June 1998.

[7] A. R. Reibman and D. Poole. Characterizing packet loss impairments in compressed video. *IEEE ICIP*, Sept 2007.

25

[8] A. R. Reibman et al. Predicting packet-loss visibility using scene characteristics. *Packet Video*, pages 308–317, Sept 2007.

30

[9] JP002006033722AA, NTT, "Image quality control method and image quality control system"

[10]

<http://www.telchemy.com/appnotes/Understanding%20IP%20Video%20Quality%20Metrics.pdf>

Claims

1. Method for assessing the quality of a video signal during encoding or compressing of
5 the video signal, the method comprising the steps of:
- a) estimating the quality, Q_{cod} , of the encoded or compressed video signal using one or more parameters; and characterized by
 - b) using the key-frame rate, kfr , of the video signal as a Group of Pictures, GOP, -
length-related parameter as at least one additional parameter to adjust the
10 estimated video signal quality, in accordance with the following equation:
$$Q_{cod} = (a1 * kfr + a2) * \exp(b*br) + c$$

wherein
a1, a2, b, c are coefficients
br is the bit-rate of the video signal.
- 15
2. The method of claim 1, wherein the one or more parameter used in step a) is selected from the set comprising: bit rate, frame rate, video resolution, codec type, content type.
3. The method of claim 1 or 2, wherein the values of the parameters used in step a) are
20 computed from the packet header information extracted from the bit-stream of the video signal and/or derived from side information.
4. The method of any of the preceding claims, wherein further additional parameter(s) used in step b) is selected from the set comprising: number of bits in an I-frame,
25 number of bits in a Group of Pictures.
5. The method of any of claims 1 to 4, wherein coefficients $a1$ and $a2$ are obtained by applying a least-square-error curve fitting procedure using the ratings of perception tests as target values.
30
6. The method of any of claims 1 to 4, wherein the coefficients $a1$ and $a2$ are dependent on additional information extracted from the bit-stream or packet headers.

7. The method of claim 6, wherein the coefficients $a1$ and $a2$ are calculated using number I of bits in an I-frame and the number G of bits in a Group of Pictures according to the following equations:

$$a1 = I / G * a1'$$

5 $a2 = I / G * a2'$

so that the estimated video signal quality is in accordance with the following equation:

$$Q_{cod} = \frac{I}{G} * (a1' * kfr * + a2') * \exp(b * br) + c$$

10 where $a1'$ and $a2'$ represent given encoder settings, wherein $a1'$ and $a2'$ are obtained by applying a least-square-error curve fitting procedure using the ratings of perception tests as target values.

8. The method of any of claims 2 to 7, wherein $fr = n/t$ is the frame rate of the video sequence, where n is the number of frames on a time window of t or more seconds.

- 15 9. The method of claim 4 to 8, wherein $kfr = fr/d$ is the key-frame-rate of the video sequence, where d is the number of frames between two I-frames of the video sequence.

- 20 10. The method of any of the preceding claims, wherein the video signal is a non-interactive streaming video or an interactive streaming video.

11. The method of any of the preceding claims, wherein the method is combined with other methods for estimating the quality impact of other impairments than encoding and compressing of the video signal, the methods comprising the steps of:

- 25 a) estimating the quality due to packet loss during transmission of the video signal,
 b) estimating the base quality of the video signal,
 c) the combination of a) and b).

12. The method of claim 11, wherein the combination is:

- 30 a) a linear function of the methods, or
 b) a multiplicative function of the methods.

13. Apparatus for assessing the quality of a video signal during encoding or compressing of the video signal, comprising:

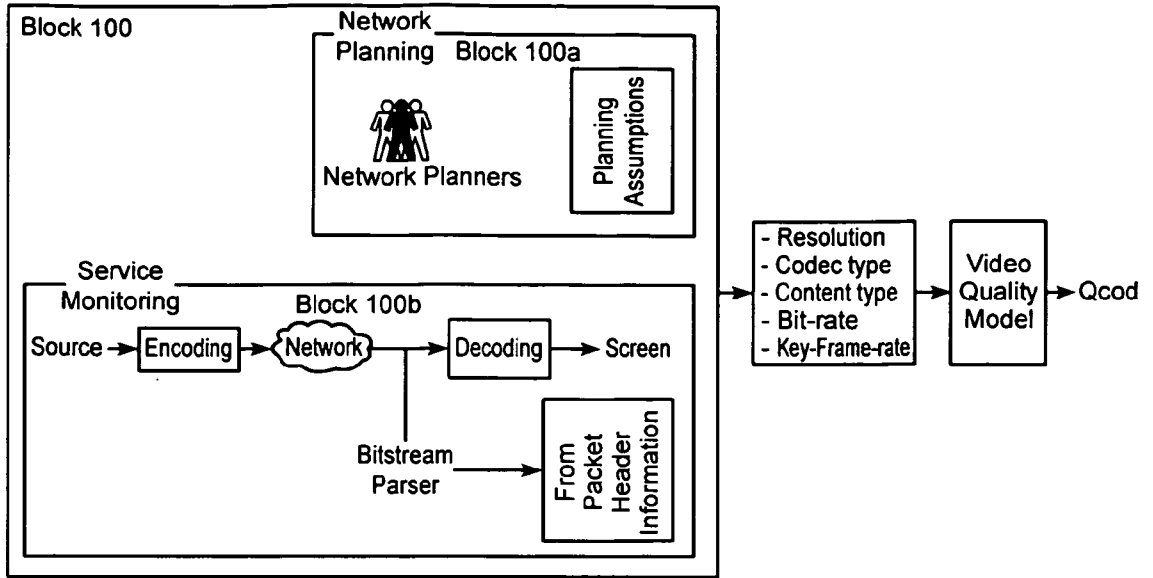
an estimator estimating the quality, Qcod, of the encoded or compressed video signal using one or more parameters; and characterized by
an adjustor using the key-frame-rate of the video signal as a Group of Pictures, GOP, -
length-related parameter as at least one additional parameter to adjust the
5 estimated video signal quality, in accordance with the following equation:

$$Q_{cod} = (a_1 * kfr + a_2) * \exp(b * br) + c$$

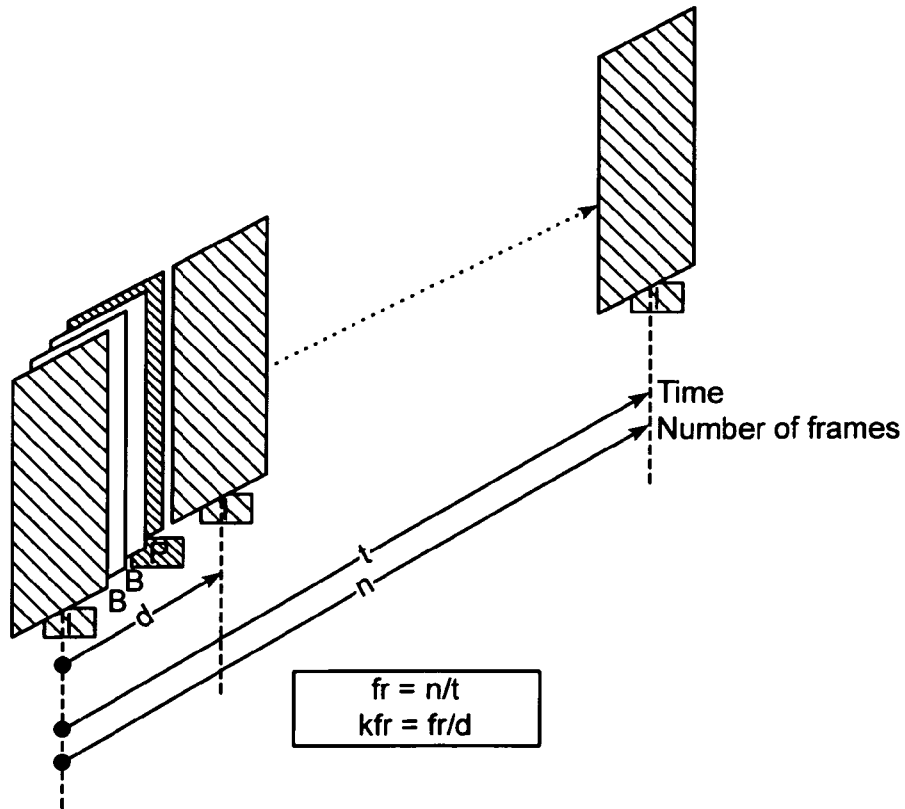
wherein

a1, a2, b, c are coefficients

br is the bit-rate of the video signal.



Figur 1



Figur 2