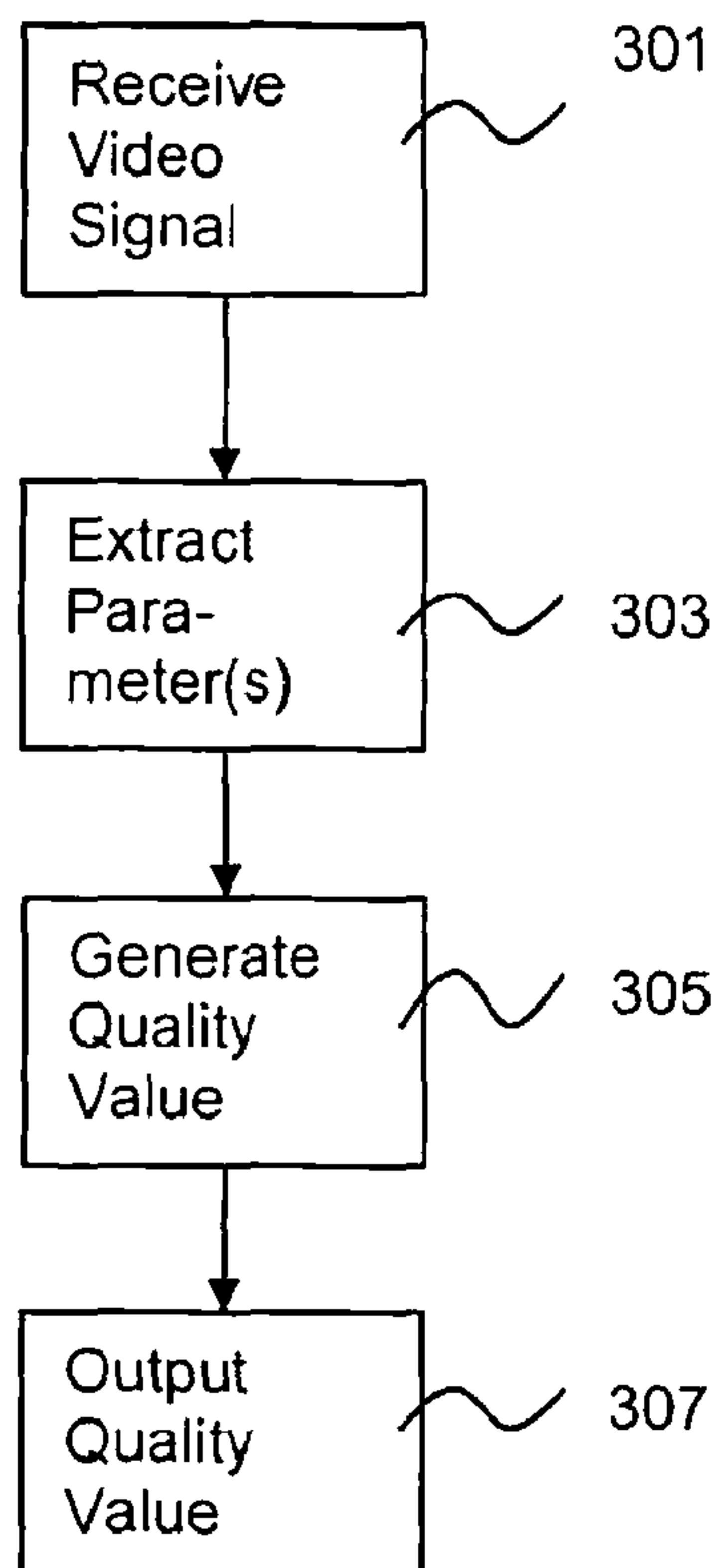




(86) Date de dépôt PCT/PCT Filing Date: 2006/10/19
(87) Date publication PCT/PCT Publication Date: 2008/04/24
(45) Date de délivrance/Issue Date: 2015/06/02
(85) Entrée phase nationale/National Entry: 2009/04/09
(86) N° demande PCT/PCT Application No.: SE 2006/050410
(87) N° publication PCT/PCT Publication No.: 2008/048155

(51) Cl.Int./Int.Cl. *H04N 21/647* (2011.01),
H04N 21/442 (2011.01)
(72) Inventeurs/Inventors:
GUSTAFSSON, JORGEN, SE;
TAN, XIANGCHUN, SE;
HEIKKILA, GUNNAR, SE;
PETTERSSON, MARTIN, SE
(73) Propriétaire/Owner:
TELEFONAKTIEBOLAGET L M ERICSSON (PUBL), SE
(74) Agent: ERICSSON CANADA PATENT GROUP

(54) Titre : PROCÉDE DE DETERMINATION D'UNE QUALITE VIDEO
(54) Title: A METHOD OF DETERMINING VIDEO QUALITY



(57) Abrégé/Abstract:

A method and a device utilizing an algorithm using measurement data derived from parameters related to a video-streaming player and/or parameters related to data transport is disclosed. The data are used as input data in a model designed to generate a value corresponding to the quality of the multimedia sequence, such as for example a MOS score.

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

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
24 April 2008 (24.04.2008)

PCT

(10) International Publication Number
WO 2008/048155 A1

(51) International Patent Classification:
H04N 7/26 (2006.01) *G06T 7/00* (2006.01)

(74) Agent: NILSSON, Charlotte; Ericsson AB, Patent Unit
3G, S-164 80 Stockholm (SE).

(21) International Application Number:
PCT/SE2006/050410

(81) Designated States (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, 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, LV, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(22) International Filing Date: 19 October 2006 (19.10.2006)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant (*for all designated States except US*): TELEFONAKTIEBOLAGET LM ERICSSON (publ)
[SE/SE]; S-164 83 Stockholm (SE).

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

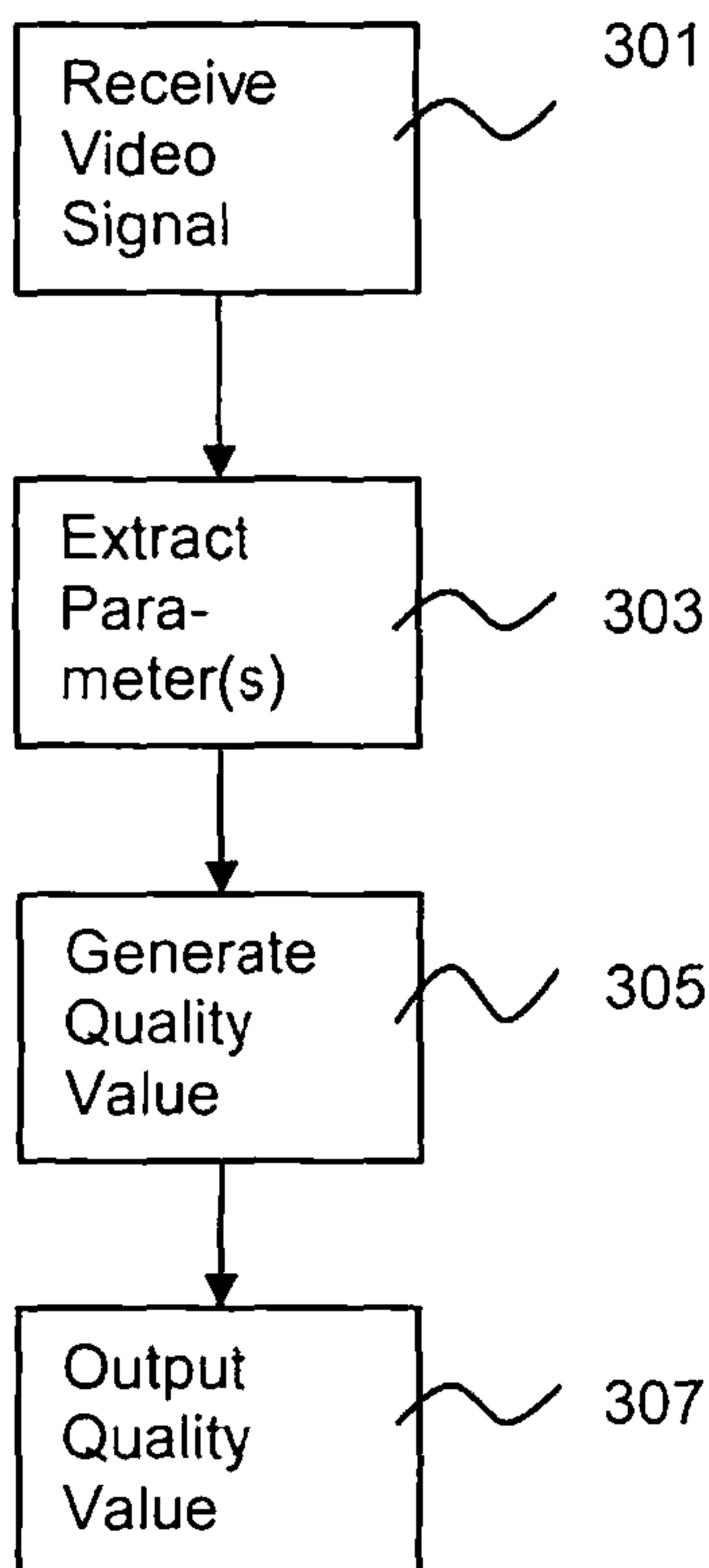
(72) Inventors; and

(75) Inventors/Applicants (*for US only*): GUSTAFSSON, Jörgen [SE/SE]; Stengatan 37, S-972 52 Luleå (SE). TAN, Xiangchun [SE/SE]; Lävågen 153, S-976 34 Luleå (SE). HEIKKILÄ, Gunnar [SE/SE]; Gäddviksvägen 31, S-954 31 Gammelstad (SE). PETTERSSON, Martin [SE/SE]; Karlsborgsvägen 135, S-195 33 Märsta (SE).

[Continued on next page]

(54) Title: A METHOD OF DETERMINING VIDEO QUALITY

(57) Abstract: A method and a device utilizing an algorithm using measurement data derived from parameters related to a video-streaming player and/or parameters related to data transport is disclosed. The data are used as input data in a model designed to generate a value corresponding to the quality of the multimedia sequence, such as for example a MOS score.



WO 2008/048155 A1

WO 2008/048155 A1



Declarations under Rule 4.17:

- *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*
- *of inventorship (Rule 4.17(iv))*

Published:

- *with international search report*

A METHOD OF DETERMINING VIDEO QUALITY

TECHNICAL FIELD

The present invention relates to a method and a device for determining the quality of a multimedia sequence.

BACKGROUND

New high performance radio networks have paved the way for new mobile services and applications. For example, many of the new services aim at enhancing the experience of a phone conversation; other services provide for transmission of video signals such as video-on-demand and other similar multimedia services. All such new services involving reception of any type of multimedia at a receiver require monitoring of the perceived received quality. This to ensure that users experience good quality and get the service they expect. In other words multimedia services such as video streaming quality perceived by the end user is one important service quality measurement for operators of all types of networks. Possible service problems need to be troubleshot.

The subjectively perceived video quality can be estimated with an objective video quality model. The video quality value can be a MOS (Mean Opinion Score) value or another suitable measure. MOS is the mean value of grades from a subjective test, where test persons grade a multimedia clip in a scale ranging from 1 for the poorest quality to 5 for the best quality.

Today, there exist a number of commercial tools for measuring and estimating multimedia quality. However, the products for determining video quality on the market today base their video quality estimation mostly on video image analysis. An approach based on video image analysis puts a high demand on the computational capacity of the tool used because the algorithms used in such analysis are in themselves very computationally demanding. This in turn makes it difficult to produce an output result in real-time.

Another problem encountered with existing solutions is that they are sensitive to change in context in the video signal because they typically base their output on an analysis between different frames of a video sequence in the case of an estimation of video quality.

Yet another problem with existing solutions, wherein original video is compared to a received video (so called full reference method), is the need for synchronization between consecutive frames in order to generate a meaningful output result.

There is therefore a need for a method that requires less computational power and which hence can be implemented in a device that is inexpensive to manufacture and easy to operate.

SUMMARY

It is an object of the present invention to overcome at least some of the problems associated with existing methods and tools for determining the quality of a multimedia sequence.

This object and others are obtained by a method and a device utilizing an algorithm using measurement data derived from input parameters related to a video-streaming player and/or parameters related to data transport. The multimedia sequence can be transmitted over any type of network, for example a radio network or a fixed network. The data are used as input data in a model designed to generate a value corresponding to the quality of the multimedia sequence, such as for example a MOS score. The method in accordance with the invention can advantageously be implemented by computer software in a computer program product.

Such an approach is significantly faster than methods based on video image analysis because the calculations required are much simpler. Typically they do not require any transformation into the frequency domain and the like.

Moreover, the exact synchronization needed for solutions requiring a reference as is the case for solutions based on a full reference video image analysis is not necessary, since a solution based on network parameters does not require a reference. Also, the calculated output, e.g. a

MOS value, is not affected if the content is switched as would be the case in a model using image analysis.

BRIEF DESCRIPTION OF THE DRAWINGS

The present invention will now be described in more detail by way of non-limiting examples and with reference to the accompanying drawings, in which:

- Fig. 1 is a general view of a system for streaming video
- Fig 2 is a closer view of some parts of the system shown in Fig.1, and
- Fig 3 is a flowchart illustrating different steps performed in a device for determining perceived multimedia quality

DETAILED DESCRIPTION

In Fig. 1, a general view of a system for measuring the received quality of a multimedia sequence in accordance with the present invention is depicted. The multimedia sequence is in the following example a video clip with associated audio. The system comprises a video streaming server 101. The video streaming server 101 is the source of a video sequence with associated audio transmitted over a network 103 as a video signal. The network 103 can for example be a radio network or any other network.

The video sequence transmitted over the network 103 is received by a mobile test system 105. The mobile test system 105 typically comprises a video streaming player 107 and a device 109 including a video quality model. The device 109 utilizes data obtained from the video streaming player 107 and from the transport layer in order to generate a value that is an estimation of a current end-user perceived video, audio or total quality.

In Fig. 2, the mobile test system 105 is shown in more detail. The test system 105 is designed to upon reception of a video sequence estimate perceived video quality for video in a computational efficient way. The estimation is done by, in real-time, collecting measurement data from a video-streaming player (video streaming), and/or from data transport, in particular IP layer and above. The data is input and combined in a video quality estimation model and a video quality value is calculated (for example a MOS score).

The calculated output (MOS value) is an estimation of end-user perceived video and audio quality. The calculated MOS value is intended to be close to the result of a subjective test, where test persons view and grade a video clip. The total streaming quality (MOS value) is expressed as a function of the input variables. Both a linear and non-linear model can be used to estimate the MOS value. Also, a combination of linear and non-linear functions may be employed.

The model is then designed to predict one or many of:

- Total quality
- Video quality
- Audio quality

The perhaps most common video quality problems for mobile video streaming are high compression (due to low bandwidth over the network), video quality degradation due to packet loss, long initial buffering and re-buffering in the middle of a video. In a preferred embodiment the model takes into account all of these quality problems and the output, for example presented as a MOS value, then reflects the perceived video quality.

The input to the model includes a number of these measurement parameters. For example one or many of the following parameters may be utilized:

- Audio Codec
- Video Codec
- Total coded bit rate
- Video bit rate
- Audio bit rate
- Video frame rate
- Packet loss rate
- Length of initial buffering (as absolute or relative percent of video clip time)
- Number of re-buffering events

- Re-buffering frequency (similar to number of re-buffering events)
- Start time of re-buffering event (as absolute time or relative percent of total clip time)
- Length of re-buffering (as absolute time or relative percent of total clip time)
- Data throughput
- IP packet jitter
- IP packet size
- RTP/RTCP signaling
- Buffer size of video client

In different scenarios different parameters may prove to be more important than others when determining a video quality value. To give an example, in some applications the parameters corresponding to, the packet loss, the length of initial buffering, and the length of re-buffering in combination with information regarding the codec and the total coded bit rate have proven to give a good estimate of the video quality value. In such a scenario only a small sub-set of parameters are necessary in order to generate a good estimate of the current quality value. This is advantageous because the model can then be made simpler and hence the implementation in a device for generating a quality value can be faster and less expensive.

If one or many of the input parameters important to the situation at hand are not available, other parameters can be used to calculate the ones missing. Thus, if for example "Total coded bit rate" is not available, "Data throughput", "IP packet size" and "Buffer size of client" can be used to estimate the total coded bit rate. Buffering input parameters, such as "number of re-bufferings", "re-buffering frequency", "start time of re-buffering", "Length of initial buffering" and "length of re-buffering", can in the same manner be estimated based on total coded bit rate, throughput and knowledge about the multimedia player's buffer size. Other relationships may be used to estimate other important parameters. A combination effect of many parameters can be used to estimate the quality. For example, the combination effect of simultaneous packet loss and re-buffering can be used. The quality is not necessary a simple addition of quality degradation for packet loss and re-buffering. In fact, in some

application the relationship is conceivably more complex and requires a more complex model in order to correctly model the perceived video quality.

The video quality model is tuned based on subjective tests, where test persons have viewed and graded video clips with long and short initial buffering and long and short middle buffering packet, various loss rates and potentially other quality degrading effects. Also the number of middle buffering may be varied to tune the model even better. Other ways to tune the model include tuning with help of other objective models, etc.

In Fig. 3 different steps performed in the video quality modeling device as described above are shown. Thus, first in a step 301 the video player receives a video signal representing a video sequence. The video sequence may or may not be associated with audio. Next, in a step 303, a number of data for different parameters relating to the received video signal and/or video player are transferred to the video quality modeling device. The data typically relate to one or many of the parameters listed above in conjunction to Fig 2, e.g. Video Codec, Audio Codec, Total coded bit rate, Video bit rate, Audio bit rate, Video frame rate, Packet loss rate, Start time of re-buffering (as absolute time or relative percent of total clip time), Length of initial buffering (as absolute or relative percent of video clip time), Number of re-buffering events, Re-buffering frequency (similar to number of re-buffering events), Length of re-buffering (as absolute time or relative percent of total clip time), Data throughput, IP packet jitter, IP packet size, RTP/RTCP signaling, Buffer size of video client or other parameters relevant for the particular situation as the situation may be.

Thereupon, in a step 305, a value indicative of the quality of the received video signal is calculated. Depending on what the device is designed to generate as an output the value can be all or a subset of total quality, video quality and audio quality. The calculations performed in step 305 can differ for different applications and scenarios and will be described more closely below. Finally, in a step 307, the device outputs a result. For example, the video quality value may be presented as a numerical value such as a MOS. This value is then used as input data for optimizing the network or similar tasks.

In accordance with one preferred embodiment, the basic function of the model can be described with:

$$TOT_MOS_{pred} = func(Qual_{encoding}, Qual_{buff}, Qual_{pl})$$

The base quality (quality of the encoding, $Qual_{encoding}$) can be described with a function:

$$y = c_0 - c_1 \cdot e^{-\lambda \cdot x}$$

where c_0 , c_1 and λ are constants. The constants have different values for different codecs.

$Qual_{buff}$ and $Qual_{pl}$ reduce the output MOS value based on initial buffering time, re-buffering percentage, re-buffering frequency and packet loss, respectively.

The packet loss effect for a logging period with packet loss can be expressed by:

$$Qual_{PL} = const * (Qual_{encoding} - 1) \cdot \xi + 1$$

where the factor ξ is defined as

$$\xi = \frac{PLR_u - PLR_{mean}}{PLR_u - PLR_l}$$

in which, PLR_u and PLR_l are the upper and lower packet loss rate limits respectively and PLR_{mean} is the average packet loss rate of the current logging window; hence $0 \leq \xi \leq 1$

$$PLR_{mean} = \frac{1}{N} \cdot \sum_{i=1}^N PLR_i$$

For packet loss a lower and upper limit can be defined, based on subjective tests. Packet loss rate (PLR) lower than the lower limit are then counted as non-visible and will hence not affect the MOS value. At the other end of the scale, a PLR equal to or above the upper limit

are per default set to be equally very bad (worst) quality. Thus, the following restrictions are applied on the instant PLR values:

$$PLR_i = \min(PLR_j, PLR_u), \text{ and}$$

$$PLR_i = \max(PLR_j, PLR_l)$$

Adding re-buffering and initial buffering effect on quality is then for example done by:

$$TOT_MOS_{pred} = Qual_{PL} - Qual_{buff}$$

and $Qual_{buff}$ is calculated by:

$$Qual_{buff} = C_0 + C_1 \cdot INIT_PERC + C_2 \cdot BUF_PERC + C_3 \cdot BUF_FRQ, \text{ and}$$

Where the variables are:

BUF_PERC: Re-buffering time / (playing time + rebuffering time + initial buffering)

INIT_PERC: Initial buffering time / (playing time + rebuffering time + initial buffering)

BUF_FRQ: Number of rebuffering events per minute

It may be that the effect of the packet loss rate, re-buffering and initial buffering is better modeled as a function:

$$TOT_MOS_{pred} = f(Qual_{PL}, Qual_{buff}),$$

rather than a pure linear effect.

To handle severe packet losses and the effect on perceived quality high packet loss rates are preferably weighted higher than lower rates. To handle long re-buffering times the output MOS score may be truncated for very long re-buffering percentages. For example, 67% or

higher re-buffering percent could be modeled to always result in the lowest video quality, e.g. MOS = 1.

Using the invention a combination of input parameters can be used to estimate the complete video quality. One advantage of the invention is that only a small subset of parameters for a video sequence are required in order to generate a good video quality value. For example, the combination of interruptions (re-buffering) and packet loss can alone be used to estimate quality in addition with codec and bit-rate.

The invention is computational efficient due to the fact that the video images are not analyzed. The input parameters are few; still good video quality estimation can be obtained. The invention also takes into account the most important video quality degradation factors: coding quality and degradation due to transport errors. Coding quality estimation uses information about codec. Transport errors are essentially low throughput or packet loss. Low throughput might cause long initial buffering and re-buffering events. Packet losses will cause image and/or audio distortions and quality degradation due to temporal problems in the video.

The invention as described herein is, due to the fact that it is computational efficient, able to produce a video quality value in real-time. It is hence very well suited for test situations where a score must be produced in the moment of measurement.

In addition using the invention as described herein there is no need for exact synchronization as is the case solutions for based on a video image analysis, since a solution based on network parameters does not require a reference. Also, the calculated output, e.g. a MOS value, is not affected if the content is switched as would be the case in a model using image analysis.

Amended page

10

CLAIMS

1. A method of determining multimedia quality, comprising the steps of:
 - receiving a multimedia sequence;
 - extracting a number of input transport layer parameters derived from transmission including the length of re-buffering and multimedia player parameters related to the received sequence, wherein said length of re-buffering is a value representing buffering subsequent to any initial buffering of the received multimedia sequence expressed as an absolute time or a value that is relative to the total time of the received multimedia sequence, and wherein said multimedia player parameters comprise a value representing a type of codec in the multimedia player, and wherein the input parameters further include the coded bitrate of the multimedia sequence;
 - generating a quality value corresponding to a subjective quality score based on at least the extracted input parameters using a quality estimation model device, wherein said multimedia quality model device is tuned to generate an output that represents an estimation of an end-user perceived quality; and
 - outputting the quality value as a measure of the quality of the received sequence.
2. The method according to claim 1, wherein the input parameters comprise parameters related to a video-streaming player.
3. The method according to claim 1 or 2, wherein at least one of the following additional parameters: packet loss, the length of initial buffering and the length of re-buffering are utilized to generate a quality value.
4. The method according to any one of claims 1 – 3, wherein the generated quality value corresponds to a Mean Opinion Score (MOS) value.
5. The method according to claim 4, wherein the generated quality value is calculated using a function of packet loss rate, re-buffering and initial buffering data.

Amended page

11

6. The method according to any one of claims 1 – 5, wherein a quality value for one or many of total quality, video quality or audio quality is generated

7. The method according to any one of claims 1 – 6, wherein the multimedia sequence is received over a radio network.

8. The method according to any one of claims 1 – 5, wherein one or several of the input parameters are derived or estimated from other parameters.

9. A device for determining multimedia quality, comprising:

-means for receiving a multimedia sequence;

-means for extracting input transport layer parameters derived from transmission including the length of re-buffering and multimedia player parameters related to the received sequence, wherein said length of re-buffering is a value representing buffering subsequent to any initial buffering of the received multimedia sequence expressed as an absolute time or a value that is relative to the total time of the received multimedia sequence, and wherein said multimedia player parameters comprise a value representing a type of codec in the multimedia player, and wherein the input parameters further include the coded bitrate of the multimedia sequence;

-means for generating a quality value corresponding to a subjective quality score based on at least the extracted input parameters, using a multimedia quality model tuned to generate an output that represents an estimation of an end-user perceived quality; and

-means for outputting the quality value as a measure of the quality of the received sequence.

10. The device according to claim 9, wherein the input parameters comprise parameters related to a video-streaming player.

11. The device according to claim 9 or 10, wherein the device is designed to utilize at least one of the following additional parameters: packet loss, the length of initial buffering, and the length of re-buffering to generate a quality value.

Amended page

12

12. The device according to any one of claims 9 – 11, wherein the generated quality value corresponds to a Mean Opinion Score (MOS) value.

13. The device according to claim 12, wherein the device is adapted to calculate the generated quality value using a function of packet loss rate, re-buffering and initial buffering data.

14. The device according to any one of claims 9 – 13, wherein the device is adapted to generate a quality value for one or many of total quality, video quality or audio quality.

15. The device according to any one of claims 9 – 14, wherein the device is portable and adapted to receive the multimedia sequence over a radio network.

16. The device according to any one of claims 9 – 15, wherein one or several of the input parameters are derived or estimated from other parameters.

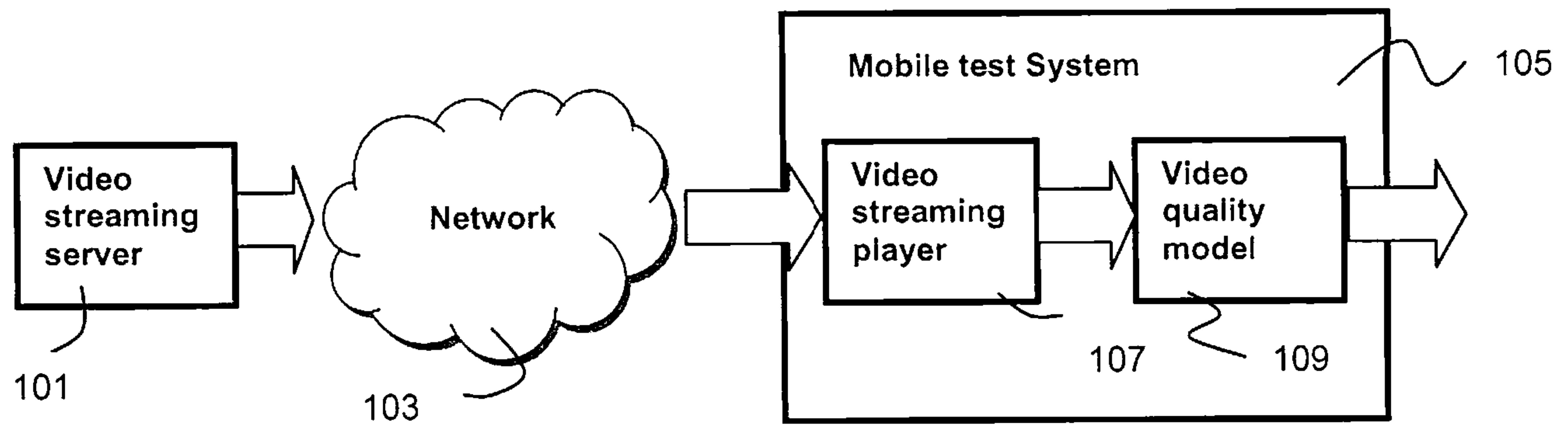


Fig.1

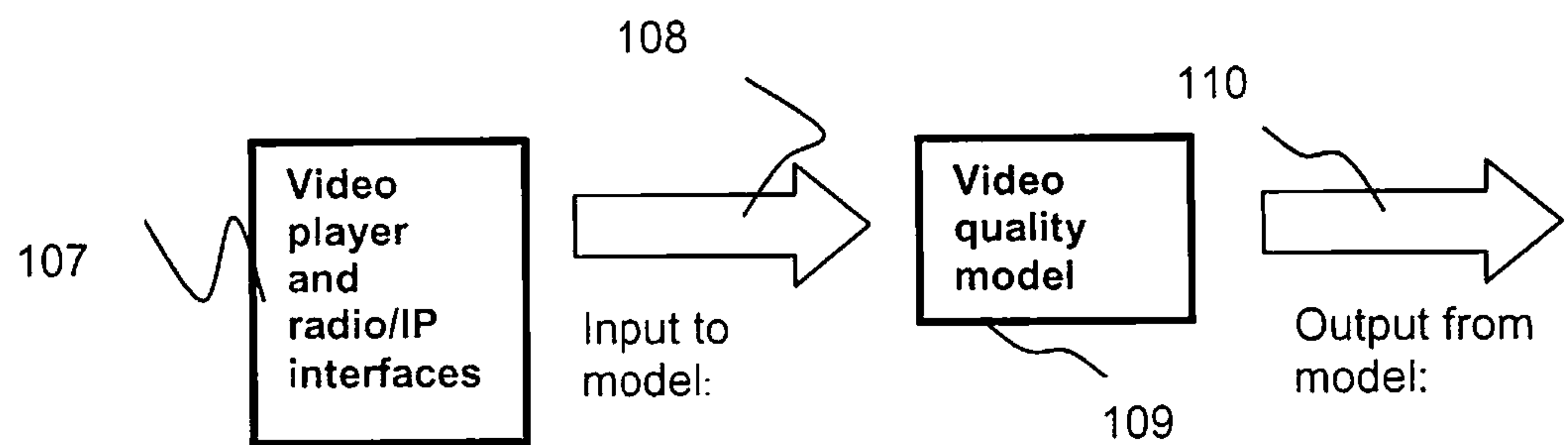


Fig.2

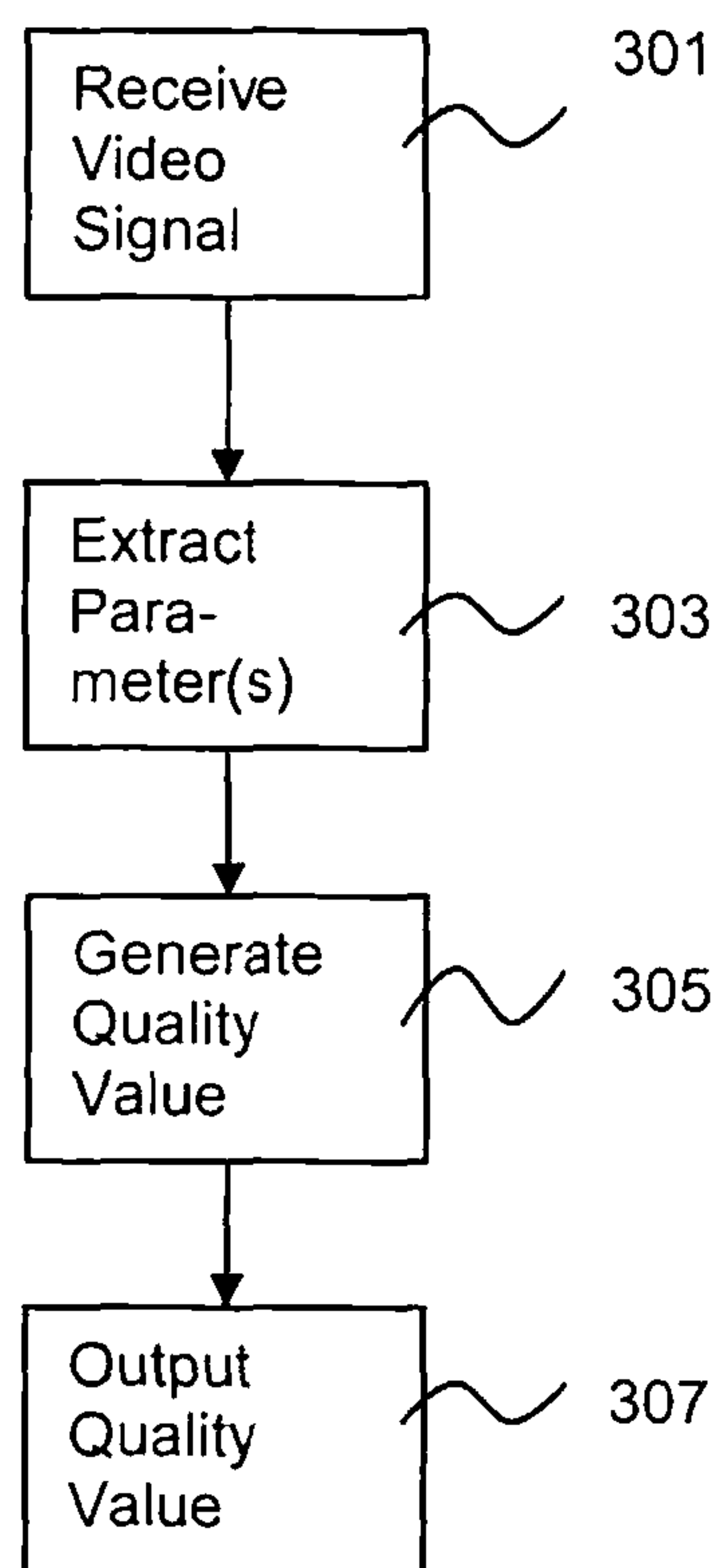


Fig.3

