

METHOD AND APPARATUS FOR EVALUATING MULTIMEDIA QUALITY

TECHNICAL FIELD

[0001] The present invention relates to the multimedia field, and in particular, to a method and an apparatus for evaluating multimedia quality.

BACKGROUND

[0002] With development of network technologies and coming of a new age of multimedia, a video on demand, a web television, videotelephony, and the like have become main services of a broadband network, and these services will also become main services of a 3rd Generation (3G) wireless network. Various multimedia processing and communications technologies emerge one after another. For a multimedia service, an amount of data is large, a requirement for real-time performance is high, and user sensitivity is strong; therefore, multimedia quality evaluation has a very important significance for a manufacturer and an operator of multimedia communications equipment. If an equipment manufacturer can provide a convincing multimedia quality evaluation result for a system, this has a very great effect of promoting sales of a product of the equipment manufacturer; and for the operator, multimedia quality evaluation data may be used for service promotion and publicity of the operator. In addition, if an automatic and real-time multimedia quality evaluation method can be developed, both the equipment manufacturer and the operator can perform real-time monitoring on a multimedia device based on this implementation, so as to help locate a problem and diagnose a fault, to ensure a user's requirement of experiencing the multimedia service.

[0003] Multimedia quality is a measurement for measuring distortion of digital multimedia relative to an original signal. Key factors that affect multimedia communication quality are video quality, audio quality, a video distortion level, and an audio distortion level. A video signal and an audio signal in a multimedia sequence need to go through phases such as sampling, quantification, compression coding, network transmission, decoding, and restoration of an analog signal, where an error and information distortion are possibly introduced in each phase, resulting in low user satisfaction. Damage and distortion of an audio or a video both cause declined multimedia experience; therefore, how to obtain the multimedia quality according to a combination of the video

quality and the audio quality becomes a key problem.

[0004] A general multimedia quality evaluation method first separately evaluates audio sequence quality and video sequence quality in the multimedia sequence on a condition that there is network damage, and then combines the audio sequence quality and the video sequence quality by using a specific polynomial formula, to obtain the multimedia quality.

[0005] When there is no packet loss on a network, both the audio and the video have relatively stable reference quality, and a feeling of a person for the multimedia quality is reference quality obtained by directly combining the audio and the video. However, when there is a packet loss, the person feels that the multimedia quality that is originally stable encounters a sudden quality deterioration, that is, the distortion of the multimedia sequence caused by the packet loss, but it is not the case that the video and the audio are, in a case of the packet loss, separately evaluated before being combined. The existing multimedia quality evaluation method cannot directly reflect impact of the packet loss on the multimedia sequence and the distortion of the multimedia sequence, and does not accord with a subjective feeling of a person.

SUMMARY

[0006] Embodiments of the present invention provide a method and an apparatus for evaluating multimedia quality, which accord with a subjective feeling of a person, and can accurately and effectively evaluate the multimedia quality.

[0007] According to one aspect, an embodiment of the present invention provides a method for evaluating multimedia quality, where the method includes: obtaining reference quality of a video and final quality of the video of a multimedia sequence, and reference quality of an audio and final quality of the audio of the multimedia sequence; determining reference quality of the multimedia sequence according to the reference quality of the video and the reference quality of the audio; determining a distortion value of the multimedia sequence according to the reference quality of the video, the final quality of the video, the reference quality of the audio, and the final quality of the audio; and determining multimedia quality of the multimedia sequence according to the reference quality of the multimedia sequence and the distortion value of the multimedia sequence.

[0008] According to another aspect, an embodiment of the present invention provides a method for evaluating multimedia quality, where the method includes: dividing a multimedia sequence into N multimedia segments, where N is a positive integer and N is greater than or equal to 2; evaluating multimedia quality of each multimedia segment of the N multimedia segments; and determining multimedia quality of the multimedia sequence according to the multimedia quality of each multimedia segment of the N multimedia segments.

[0009] According to still another aspect, an embodiment of the present invention provides an apparatus for evaluating multimedia quality, where the apparatus includes: a first obtaining module, configured to obtain reference quality of a video and final quality of the video of a multimedia sequence, and reference quality of an audio and final quality of the audio of the multimedia sequence; a reference quality determining module, configured to determine reference quality of the multimedia sequence according to the reference quality of the video and the reference quality of the audio that are obtained by the first obtaining module; a distortion value determining module, configured to determine a distortion value of the multimedia sequence according to the reference quality of the video, the final quality of the video, the reference quality of the audio, and the final quality of the audio that are obtained by the first obtaining module; and a multimedia quality determining module, configured to determine multimedia quality of the multimedia sequence according to the reference quality of the multimedia sequence that is determined by the reference quality determining module and the distortion value of the multimedia sequence that is determined by the distortion value determining module.

[0010] According to still another aspect, an embodiment of the present invention provides an apparatus for evaluating multimedia quality, where the apparatus includes: a segmenting module, configured to divide a multimedia sequence into N multimedia segments, where N is a positive integer and N is greater than or equal to 2; an evaluating module, configured to evaluate multimedia quality of each multimedia segment of the N multimedia segments; and a processing module, configured to determine multimedia quality of the multimedia sequence according to the multimedia quality of each multimedia segment of the N multimedia segments.

[0011] Based on the foregoing technical solutions, the method and the apparatus for evaluating multimedia quality according to the embodiments of the present invention, by determining the multimedia quality of the multimedia sequence according to the reference quality of the multimedia sequence and the distortion value of the multimedia sequence, can directly reflect the distortion of the multimedia sequence, accord with a subjective feeling of a person, and therefore can accurately and effectively evaluate the multimedia quality.

BRIEF DESCRIPTION OF DRAWINGS

[0012] To describe the technical solutions in the embodiments of the present invention more clearly, the following briefly introduces the accompanying drawings required for describing the embodiments of the present invention. Apparently, the accompanying drawings in the following description show merely some embodiments of the present invention, and a person of ordinary skill in the art may still derive other drawings from these accompanying drawings without creative

efforts.

[0013] FIG. 1 is a schematic flowchart of a method for evaluating multimedia quality according to an embodiment of the present invention;

[0014] FIG. 2 is a schematic flowchart of a method for determining a distortion value of a multimedia sequence according to an embodiment of the present invention;

[0015] FIG. 3 is another schematic flowchart of a method for evaluating multimedia quality according to an embodiment of the present invention;

[0016] FIG. 4 is still another schematic flowchart of a method for evaluating multimedia quality according to an embodiment of the present invention;

[0017] FIG. 5 is a schematic flowchart of a method for evaluating multimedia quality according to another embodiment of the present invention;

[0018] FIG. 6 is a schematic block diagram of an apparatus for evaluating multimedia quality according to an embodiment of the present invention;

[0019] FIG. 7 is a schematic block diagram of a distortion value determining module according to an embodiment of the present invention;

[0020] FIG. 8 is another schematic block diagram of an apparatus for evaluating multimedia quality according to an embodiment of the present invention; and

[0021] FIG. 9 is a schematic block diagram of an apparatus for evaluating multimedia quality according to another embodiment of the present invention.

DESCRIPTION OF EMBODIMENTS

[0022] The following clearly and completely describes the technical solutions in the embodiments of the present invention with reference to the accompanying drawings in the embodiments of the present invention. Apparently, the described embodiments are a part rather than all of the embodiments of the present invention. All other embodiments obtained by a person of ordinary skill in the art based on the embodiments of the present invention without creative efforts shall fall within the protection scope of the present invention.

[0023] FIG. 1 is a schematic flowchart of a method 100 for evaluating multimedia quality according to an embodiment of the present invention. As shown in FIG. 1, the method 100 includes:

[0024] S110: Obtain reference quality of a video and final quality of the video of a multimedia sequence, and reference quality of an audio and final quality of the audio of the multimedia sequence.

[0025] S120: Determine reference quality of the multimedia sequence according to the reference quality of the video and the reference quality of the audio.

[0026] S130: Determine a distortion value of the multimedia sequence according to the reference quality of the video, the final quality of the video, the reference quality of the audio, and the final quality of the audio.

[0027] S140: Determine multimedia quality of the multimedia sequence according to the reference quality of the multimedia sequence and the distortion value of the multimedia sequence.

[0028] A feeling of a person for multimedia quality is that distortion is superposed on stable multimedia reference quality. What the multimedia reference quality describes is multimedia compression quality that is caused by coding compression of the video or the audio in multimedia. Due to a factor such as a packet loss or multimedia freezing, the person feels the distortion of the multimedia sequence. The distortion is a relative deterioration level that is relative to the reference quality, that is, the multimedia quality deteriorates from relatively stable reference quality. In this embodiment of the present invention, after the reference quality of the video and the final quality of the video of the multimedia sequence, and the reference quality of the audio and the final quality of the audio of the multimedia sequence are obtained, the reference quality of the multimedia sequence is determined according to the reference quality of the video and the reference quality of the audio, the distortion value of the multimedia sequence is determined according to the reference quality of the video, the final quality of the video, the reference quality of the audio, and the final quality of the audio, and finally the multimedia quality of the multimedia sequence is determined according to the reference quality of the multimedia sequence and the distortion value of the multimedia sequence. In this way, the distortion of the multimedia sequence is comprehensively understood, but is not a combination after the video and the audio are separately understood.

[0029] Therefore, the method for evaluating multimedia quality according to this embodiment of the present invention, by determining the multimedia quality of the multimedia sequence according to the reference quality of the multimedia sequence and the distortion value of the multimedia sequence, can directly reflect the distortion of the multimedia sequence, accords with a subjective feeling of a person, and therefore can accurately and effectively evaluate the multimedia quality.

[0030] In S110, the reference quality of the video, the final quality of the video, the reference quality of the audio, and the final quality of the audio of the multimedia sequence are obtained. The reference quality of the video is video compression quality that is caused due to the coding compression of the video, and specifically may be obtained through calculation according to coding bit rates of videos with different coding types and resolutions; and the reference quality of the audio is audio compression quality that is caused due to the coding compression of the audio, and specifically may be obtained through calculation according to coding bit rates of audios with

different coding types, sampling frequencies, the number of channels, and so on. The final quality of the video and the final quality of the audio are the obtained final experience quality of the video and the audio of the multimedia sequence, where network damage (for example, the packet loss, a jitter, and so on) separately occurs in the video and the audio, and specifically may be obtained through calculation by performing an audio and video analysis on the multimedia sequence.

[0031] In S120, the reference quality of the multimedia sequence is determined according to the reference quality of the video and the reference quality of the audio. This embodiment of the present invention does not limit a specific manner in which the reference quality of the multimedia sequence is determined. For example, reference quality Q_{av}' of the multimedia sequence may be determined by the following equation (1):

$$Q_{av}' = a_1 \cdot Q_{v_coding} + a_2 \cdot Q_{a_coding} + a_3 \cdot Q_{v_coding} \cdot Q_{a_coding} + a_4 \quad (1)$$

where, Q_{v_coding} and Q_{a_coding} are respectively the reference quality of the video and the reference quality of the audio, a_1, a_2, a_3 and a_4 are parameters related to a space resolution and a display mode of the video, and their values are obtained by using data training, and are basically all decimal fractions between 0 and 1, for example, when the resolution is (128×96), $a_1=0.207962, a_2=0.124365, a_3=0.179018, a_4=0.5456$.

[0032] In S130, the distortion value of the multimedia sequence is determined according to the reference quality of the video, the final quality of the video, the reference quality of the audio, and the final quality of the audio. As shown in FIG. 2, S130 also includes:

[0033] S131: Determine a distortion value of the video according to the reference quality of the video and the final quality of the video.

[0034] S132: Determine a distortion value of the audio according to the reference quality of the audio and the final quality of the audio.

[0035] S133: Determine the distortion value of the multimedia sequence according to the distortion value of the video and the distortion value of the audio.

[0036] The distortion value of the video may be obtained by subtracting the final quality of the video from the reference quality of the video, and the distortion value of the audio may be obtained by subtracting the final quality of the audio from the reference quality of the audio.

[0037] In S133, the distortion value of the multimedia sequence is determined according to the distortion value of the video and the distortion value of the audio. Specifically, the distortion value of the multimedia sequence may be obtained in the following manner.

[0038] A distortion factor of the video is determined according to the reference quality of the video and the distortion value of the video. A distortion factor d_v of the video is a proportion of a

distortion value D_v of the video to reference quality Q_{v_coding} of the video, and optionally may be determined by the following equation (2):

$$d_v = \frac{D_v}{Q_{v_coding}} \quad (2)$$

[0039] It should be understood that, the distortion value D_v of the video is obtained by subtracting the final quality of the video from the reference quality of the video; therefore, optionally, the distortion factor d_v of the video may also be determined by the following equation (3):

$$d_v = \frac{Q_{v_coding} - Q_v}{Q_{v_coding}} \quad (3)$$

[0040] A distortion factor of the audio is determined according to the reference quality of the audio and the distortion value of the audio. A distortion factor d_a of the audio is a proportion of a distortion value D_a of the audio to reference quality Q_{a_coding} of the audio, and optionally may be determined by the following equation (4):

$$d_a = \frac{D_a}{Q_{a_coding}} \quad (4)$$

[0041] Similarly, the distortion value D_a of the audio is obtained by subtracting the final quality of the audio from the reference quality of the audio; therefore, optionally, the distortion factor d_a of the audio may also be determined by the following equation (5):

$$d_a = \frac{Q_{a_coding} - Q_a}{Q_{a_coding}} \quad (5)$$

[0042] The distortion factor of the multimedia sequence is determined according to the distortion factor of the video and the distortion factor of the audio. A distortion factor d_{av} of the multimedia sequence is determined by the distortion factor d_v of the video and the distortion factor d_a of the audio. An increase of the distortion factor d_v of the video or the distortion factor d_a of the audio increases the distortion factor d_{av} of the multimedia sequence, and their relationship may be linear, may also be non-linear, and may also be a combination of linear and non-linear. This embodiment of the present invention does not limit a specific manner in which the distortion factor d_{av} of the multimedia sequence is calculated according to the distortion factor d_v of the video and the distortion factor d_a of the audio, for example, d_{av} may be determined by the following equation (6) or (7):

$$d_{av} = \frac{a_5 \cdot d_v + a_6 \cdot d_a}{1 + a_5 \cdot d_v + a_6 \cdot d_a} \quad (6)$$

$$d_{av} = a_5 + a_6 \cdot d_v + a_7 \cdot d_a \quad (7)$$

where, a_5, a_6, a_7 are constants, and their values are related to a coding type and a video resolution and meet a condition that d_{av} increases when d_v or d_a increases; and specific numerical values may be obtained by using an experiment.

[0043] The distortion value of the multimedia sequence is determined according to the reference quality of the multimedia sequence and the distortion factor of the multimedia sequence. For example, a distortion value D_{av} of the multimedia sequence may be determined by the following equation (8):

$$D_{av} = (Q_{av}' - Q_{\min}) \cdot d_{av} \quad (8)$$

where, $Q_{\min}$ is a constant, and indicates minimum multimedia quality, for example, when a rating is a 5-point system, the minimum quality is 1.

[0044] In S140, the multimedia quality of the multimedia sequence is determined according to the reference quality of the multimedia sequence and the distortion value of the multimedia sequence. After the reference quality Q_{av}' of the multimedia sequence and the distortion value D_{av} of the multimedia sequence are determined, multimedia quality Q_{av} of the multimedia sequence is obtained according to Q_{av}' and D_{av} , and their relationship may be indicated by the following equation (9):

$$Q_{av} = f(Q_{av}', D_{av}) \quad (9)$$

[0045] For example, the relationship may be expressed as the following equation (10):

$$Q_{av} = Q_{av}' - D_{av} \quad (10)$$

[0046] The formulas (9) and (10) indicate that the multimedia quality of the multimedia sequence is a result obtained after the distortion is superposed on the reference quality of the multimedia sequence, and the feeling of the person for the multimedia quality is exactly that the distortion is superposed on stable multimedia reference quality. Therefore, the method for evaluating multimedia quality according to this embodiment of the present invention accords with a cognition characteristic of the person.

[0047] In this way, the method for evaluating multimedia quality according to this embodiment of the present invention, by determining the multimedia quality of the multimedia sequence according to the reference quality of the multimedia sequence and the distortion value of the multimedia sequence, can directly reflect the distortion of the multimedia sequence, accords with

the subjective feeling of the person, and therefore can accurately and effectively evaluate the multimedia quality.

[0048] Optionally, the multimedia quality of the multimedia sequence may be obtained after the equations (1), (3), (5), (6), and (8) are substituted into the equation (10). In this case, the multimedia quality Q_{av} of the multimedia sequence is a function of the reference quality Q_{v_coding} of the video, the reference quality Q_{a_coding} of the audio, final quality Q_v of the video, and final quality Q_a of the audio.

[0049] Optionally, for the equation (1), the following equation (11) may be obtained by substituting parameters of the equation (1) and performing arrangement:

$$\begin{aligned}
 Q_{av}' &= a_1 \cdot (Q_v + D_v) + a_2 \cdot (Q_a + D_a) + a_3 \cdot (Q_v + D_v) \cdot (Q_a + D_a) + a_4 \\
 &= a_1 \cdot Q_v + a_1 \cdot D_v + a_2 \cdot Q_a + a_2 \cdot D_a + a_3 \cdot (Q_v \cdot Q_a + Q_v \cdot D_a + D_v \cdot Q_a + D_v \cdot D_a) + a_4 \\
 &= (a_1 \cdot Q_v + a_2 \cdot Q_a + a_3 \cdot Q_v \cdot Q_a + a_4) + (a_1 \cdot D_v + a_2 \cdot Q_a + a_3 \cdot (Q_v \cdot D_a + D_v \cdot Q_a + D_v \cdot D_a)) \\
 &= Q'' + f_1(Q_v, Q_a, D_v, D_a)
 \end{aligned}
 \tag{11}$$

[0050] In the equation (11), Q'' indicates $a_1 \cdot Q_v + a_2 \cdot Q_a + a_3 \cdot Q_v \cdot Q_a + a_4$, that is, one quality score is obtained according to the final quality of the video and the final quality of the audio. $a_1 \cdot D_v + a_2 \cdot Q_a + a_3 \cdot (Q_v \cdot D_a + D_v \cdot Q_a + D_v \cdot D_a)$ is expressed in a form of a function $f_1(Q_v, Q_a, D_v, D_a)$, and then the equation (10) for calculating the multimedia quality may be expressed as the following equation (12):

$$Q_{av} = (Q'' + f_1(Q_v, Q_a, D_v, D_a)) - (Q'' + f_1(Q_v, Q_a, D_v, D_a) - Q_{\min}) \cdot d_{av} \tag{12}$$

where, a multimedia distortion factor d_{av} may be expressed in multiple forms of functions, and therefore may be expressed in a form of $f_2(Q_v, Q_a, D_v, D_a)$, and then the equation (12) may be expanded as the following equation (13):

$$\begin{aligned}
 Q_{av} &= Q'' \cdot (1 - f_2(Q_v, Q_a, D_v, D_a)) + \\
 &\quad f_1(Q_v, Q_a, D_v, D_a) \cdot (1 - f_2(Q_v, Q_a, D_v, D_a)) + Q_{\min} \cdot f_2(Q_v, Q_a, D_v, D_a) \\
 &= Q'' \cdot f_3(Q_v, Q_a, D_v, D_a) + f_4(Q_{\min}, Q_v, Q_a, D_v, D_a)
 \end{aligned}
 \tag{13}$$

[0051] $1 - f_2(Q_v, Q_a, D_v, D_a)$ is expressed by using a function $f_3(Q_v, Q_a, D_v, D_a)$, $f_1(Q_v, Q_a, D_v, D_a) \cdot (1 - f_2(Q_v, Q_a, D_v, D_a)) + Q_{\min} \cdot f_2(Q_v, Q_a, D_v, D_a)$ is expressed by using a function $f_4(Q_{\min}, Q_v, Q_a, D_v, D_a)$, and then a calculation method for multimedia quality may be expanded as another expression form described above, that is, the multimedia quality Q_{av} is a function of the final quality Q_v of the video, the final quality Q_a of the audio, the distortion

value D_v of the video, and the distortion value D_a of the audio.

[0052] The method for evaluating multimedia quality according to this embodiment of the present invention, by determining the multimedia quality of the multimedia sequence according to the reference quality and the distortion value of the multimedia sequence, can directly reflect the distortion of the multimedia sequence, accords with the subjective feeling of the person, and therefore can accurately and effectively evaluate the multimedia quality.

[0053] When the audio and the video are asynchronous, the multimedia quality of the multimedia sequence possibly deteriorates. In this scenario, a problem that the audio and the video are asynchronous needs to be considered. Therefore, as shown in FIG. 3, the method 100 for evaluating multimedia quality according to this embodiment of the present invention further includes:

[0054] S150: Obtain an effect factor of audio and video asynchronization of the multimedia sequence.

[0055] S140 further includes:

[0056] S141: Determine the multimedia quality of the multimedia sequence according to the reference quality of the multimedia sequence, the distortion value of the multimedia sequence, and the effect factor of audio and video asynchronization of the multimedia sequence.

[0057] In S150, the effect factor of audio and video asynchronization of the multimedia sequence is obtained. The effect factor of audio and video asynchronization of the multimedia sequence is a function of a time difference ΔT_{syn} of audio and video asynchronization, and reflects impact of audio and video asynchronization on the multimedia quality.

[0058] In S141, the multimedia quality of the multimedia sequence is determined according to the reference quality of the multimedia sequence, the distortion value of the multimedia sequence, and the effect factor of audio and video asynchronization of the multimedia sequence. Their relationship may be expressed as the following equation (14):

$$Q_{av} = f(Q_{av}', D_{av}, f_5(\Delta T_{syn})) \quad (14)$$

where, $f_5(\Delta T_{syn})$ is the effect factor of audio and video asynchronization of the multimedia sequence.

[0059] For example, optionally, the multimedia quality Q_{av} of the multimedia sequence may be determined by the following equation (15):

$$Q_{av} = Q_{av}' - D_{av} \cdot f_5(\Delta T_{syn}) \quad (15)$$

where, $f_5(\Delta T_{syn})$ is greater than 1, and a larger $|\Delta T_{syn}|$ indicates a larger $f_5(\Delta T_{syn})$,

which makes Q_{av} smaller; and $f_5(\Delta T_{syn})$ is not limited in form, may be linear, may also be non-linear, and may also be a combination of a linear formula and a non-linear formula; for example, its specific form may be expressed as the following equation (16):

$$\begin{aligned} f_5(\Delta T_{syn}) &= b_1 \cdot |\Delta T_{syn}| + c_1 (b_1 > 0) \quad \text{or} \\ f_5(\Delta T_{syn}) &= b_1 \cdot (|\Delta T_{syn}|)^2 + c_1 \cdot |\Delta T_{syn}| + d_1 (b_1 > 0, c_1 \geq 0) \quad \text{or} \\ f_5(\Delta T_{syn}) &= b_1^{|\Delta T_{syn}|} (b_1 > 1) \end{aligned} \quad (16)$$

5 where, b_1, c_1, d_1 are constants.

[0060] Optionally, the multimedia quality Q_{av} of the multimedia sequence may also be determined by the following equation (17):

$$Q_{av} = (Q_{av}' - D_{av}) \cdot f_6(\Delta T_{syn}) \quad (17)$$

10 where, $f_6(\Delta T_{syn})$ is the effect factor of audio and video asynchronization of the multimedia sequence, its value is less than 1 and greater than 0, and a larger $|\Delta T_{syn}|$ indicates a smaller $f_6(\Delta T_{syn})$, which makes Q_{av} smaller; and $f_6(\Delta T_{syn})$ is not limited in form, may be linear, may also be non-linear, and may also be a combination of a linear formula and a non-linear formula; for example, its specific form may be expressed as the following equation (18):

$$\begin{aligned} f_6(\Delta T_{syn}) &= b_2 \cdot |\Delta T_{syn}| + c_2 (b_2 < 0) \quad \text{or} \\ f_6(\Delta T_{syn}) &= b_2 \cdot (|\Delta T_{syn}|)^2 + c_2 \cdot |\Delta T_{syn}| + d_2 (b_2 < 0, c_2 \leq 0) \quad \text{or} \\ f_6(\Delta T_{syn}) &= b_2^{|\Delta T_{syn}|} (0 < b_2 < 1) \end{aligned} \quad (18)$$

15 where, b_2, c_2, d_2 are constants.

[0061] The formulas (14), (15), and (17) indicate that the multimedia quality of the multimedia sequence is a result obtained after the distortion and the impact of audio and video asynchronization are superposed on the reference quality of the multimedia sequence, and accords with the cognition characteristic of the person.

20 **[0062]** Therefore, the method for evaluating multimedia quality according to this embodiment of the present invention, by determining the multimedia quality according to the reference quality of the multimedia sequence, the distortion value of the multimedia sequence, and the effect factor of audio and video asynchronization of the multimedia sequence, accords with the subjective feeling of the person, and therefore can accurately and effectively evaluate the multimedia quality.

25 **[0063]** With reference to FIG. 4, the following describes in detail the method for evaluating multimedia quality according to this embodiment of the present invention.

[0064] S410: Calculate reference quality of a multimedia sequence according to reference quality of a video and reference quality of an audio. For example, the foregoing equation (1) may be used.

[0065] S420: Calculate a distortion factor of the video according to the reference quality of the video and final quality of the video. For example, the foregoing equation (3) may be used.

[0066] S430: Calculate a distortion factor of the audio according to the reference quality of the audio and final quality of the audio. For example, the foregoing equation (5) may be used.

[0067] S440: Calculate a distortion factor of the multimedia sequence according to the distortion factor of the video and the distortion factor of the audio. For example, the foregoing equation (6) or (7) may be used.

[0068] S450: Calculate multimedia quality of the multimedia sequence. When audio and video asynchronization does not occur, the multimedia quality of the multimedia sequence is calculated according to the reference quality of the multimedia sequence and the distortion factor of the multimedia sequence, for example, the foregoing equations (8) and (10) may be used; and when the audio and the video are asynchronous, the multimedia quality of the multimedia sequence is calculated according to the reference quality of the multimedia sequence, the distortion factor of the multimedia sequence, and an effect factor of audio and video asynchronization of the multimedia sequence, for example, the foregoing equation (8) and equation (15) or (17) may be used.

[0069] Therefore, the method for evaluating multimedia quality according to this embodiment of the present invention, by determining the multimedia quality of the multimedia sequence according to the reference quality of the multimedia sequence and the distortion factor of the multimedia sequence, accords with the subjective feeling of the person, and can accurately and effectively evaluate the multimedia quality.

[0070] FIG. 5 is a schematic flowchart of a method 500 for evaluating multimedia quality according to an embodiment of the present invention. As shown in FIG. 5, the method 500 includes:

[0071] S510: Divide a multimedia sequence into N multimedia segments, where N is a positive integer and N is greater than or equal to 2.

[0072] S520: Evaluate multimedia quality of each multimedia segment of the N multimedia segments.

[0073] S530: Determine multimedia quality of the multimedia sequence according to the multimedia quality of each multimedia segment of the N multimedia segments.

[0074] For one multimedia sequence, due to a short-term memory of a person, the multimedia quality is understood with a deeper memory within a shorter period of time from the present time. In this embodiment of the present invention, the multimedia sequence is first divided into multiple

multimedia segments, the multimedia quality of each multimedia segment is then separately evaluated, and then the multimedia quality of the multimedia sequence is determined according to the multimedia quality of each multimedia segment.

[0075] Therefore, the method for evaluating multimedia quality according to this embodiment of the present invention, by dividing the multimedia sequence into segments, and determining the multimedia quality of the multimedia sequence according to the multimedia quality of each multimedia segment, is convenient for determining the multimedia quality of the multimedia sequence according to a level of attention that is paid by the person to each multimedia segment, accords with a cognition characteristic of the person, and can increase an accuracy of evaluating the multimedia quality.

[0076] In S510, the multimedia sequence is divided into the N multimedia segments. N is a positive integer, and N is greater than or equal to 2, that is, the multimedia sequence is divided into at least two multimedia segments.

[0077] Optionally, the multimedia sequence may be divided into N multimedia segments according to a duration. For example, starting from a first frame of the multimedia sequence, the multimedia sequence within each L-second duration is divided into one segment, where a value of L may be adjusted according to a specific situation.

[0078] Optionally, the multimedia sequence may also be divided into N multimedia segments according to a level of multimedia quality, video quality, or audio quality. People are prone to pay attention to relatively special content in the multimedia sequence, for example, a video picture affected by a bit error, a sound of a distorted audio, and so on. Therefore, multimedia segmentation takes low-quality multimedia content as a center, and divides the entire multimedia sequence into consecutive multimedia segments with the duration being about L seconds, where the value of L may be adjusted according to the specific situation. For example, the segmentation may be performed according to the following manner:

[0079] 1. Starting from the first frame, slide a window of $0.8 \times L$ seconds, and calculate the multimedia quality within each window.

[0080] 2. Select a window position with worst quality to determine one multimedia segment.

[0081] 3. Exclude the selected multimedia segment, repeat steps 1 and 2 to obtain a new multimedia segment, where an interval is controlled to be $0-0.4 \times L$ seconds if the new multimedia segment is adjacent to the determined multimedia segment.

[0082] 4. Repeat step 3 to divide the entire multimedia sequence.

[0083] 5. Combine a multimedia frame at the interval of the multimedia segments into an adjacent multimedia segment by adopting a manner of equal allocation, to implement complete

segmentation of the entire multimedia sequence.

[0084] It should be understood that, the foregoing manner of the segmentation performed according to the multimedia quality may also be replaced with segmentation performed according to the video quality or the audio quality. During the segmentation performed according to the video quality, a length unit may be a GOP (Group of Pictures, group of pictures) of the video, that is, one or more GOP lengths are used as one multimedia segment.

[0085] It should be further understood that, the manner of performing segmentation on the multimedia sequence may also be another manner; for example, the length of each segment may be different; and for another example, during the segmentation performed according to the multimedia quality, a multimedia segment with high quality may be first selected, and so on. This embodiment of the present invention does not limit a specific manner in which the multimedia sequence is divided into N multimedia segments.

[0086] In S520, the multimedia quality of each multimedia segment of the N multimedia segments is evaluated. For each multimedia segment, the multimedia quality of the multimedia segment may be obtained through calculation according to a combination of quality and distortion conditions of the audio and the video of each multimedia segment. The embodiment of the present invention does not limit a manner in which the multimedia quality of each multimedia segment is evaluated. For a specific manner, the prior art may be used, and the foregoing method 100 for evaluating multimedia quality according to the embodiment of the present invention may also be used. Details are not provided herein again.

[0087] In S530, the multimedia quality of the multimedia sequence is determined according to the multimedia quality of each multimedia segment of the N multimedia segments. After the multimedia quality of each multimedia segment of the N multimedia segments is evaluated, the multimedia quality of the multimedia sequence is obtained according to multimedia quality of the N multimedia segments.

[0088] Optionally, the multimedia quality of the multimedia sequence may be determined by performing weighted averaging on the multimedia quality of the N multimedia segments. Weighted averaging may be performed on the multimedia quality of the N multimedia segments based on an equal weight value, and weighted averaging may also be performed on the multimedia quality of the N multimedia segments based on a weight value related to the multimedia quality of each multimedia segment of the N multimedia segments, for example, lower multimedia quality of a multimedia segment indicates a higher weight. For example, quality of the multimedia sequence may be determined by using the following equation (19):

$$Q = \frac{\sum_{m \in \text{sequence}} Q_{av,m} \cdot W_m}{\sum_{m \in \text{sequence}} W_m} \quad (19)$$

where, m indicates an m^{th} multimedia segment in the multimedia sequence, $Q_{av,m}$ is multimedia quality of the m^{th} multimedia segment, W_m is its weight value, and may be an equal constant, or a weight that is applied according to a level of the multimedia quality.

5 **[0089]** Because people are easy to forget things, a multimedia segment that is seen lately is impressive, and a memory about a multimedia segment that is seen early is relatively fuzzy. Therefore, weighted averaging may also be performed on the multimedia quality of the N multimedia segments based on a weight value related to a time of each multimedia segment of the N multimedia segments, for example, a shorter temporal distance between a time of the multimedia
10 segment and a current rating time indicates a larger weight value. For example, quality of the multimedia sequence may be determined by using the following equation (20):

$$Q = \frac{\sum_{m \in \text{sequence}} Q_{av,m} \cdot W_{t_m}}{\sum_{m \in \text{sequence}} W_{t_m}} \quad (20)$$

where, t_m is a temporal distance between a time of the m^{th} multimedia segment and the current rating time, and W_{t_m} is a weight value related to this temporal distance.

15 **[0090]** Weighted averaging may also be performed on the multimedia quality of the N multimedia segments based on a weight value related to the multimedia quality and a time of each multimedia segment of the N multimedia segments. For example, quality of the multimedia sequence may be determined by using the following equation (21):

$$Q = \frac{\sum_{m \in \text{sequence}} Q_{av,m} \cdot W_m \cdot W_{t_m}}{\sum_{m \in \text{sequence}} W_m \cdot W_{t_m}} \quad (21)$$

20 where, W_m is the weight value related to the multimedia quality, W_{t_m} is the weight value related to the temporal distance, and the two weight values may also be combined into one weight value, which is not only related to the multimedia quality, but also related to the temporal distance.

[0091] Eyes of a person easily notice a multimedia segment with worse quality, and the worse
25 multimedia segment can reflect quality of the entire multimedia sequence. Therefore, optionally, K multimedia segments with worst multimedia quality may be selected from the N multimedia segments, and the multimedia quality of the multimedia sequence is determined by performing

weighted averaging on multimedia quality of the K multimedia segments, where K is a positive integer, and K is greater than or equal to 1 and less than N, and a specific value of K may be set according to an actual application scenario. For a manner in which weighted averaging is performed on the multimedia quality of the K multimedia segments, refer to the foregoing manner in which weighted averaging is performed on the multimedia quality of the N multimedia segments, that is, weighted averaging may be performed on the multimedia quality of the K multimedia segments based on an equal weight value, or based on a weight value related to the multimedia quality and/or a time of each multimedia segment of the K multimedia segments.

[0092] Optionally, if weighted averaging is performed based on the equal weight value, the quality of the multimedia sequence may be determined by using the following equation (22):

$$Q = \frac{\sum_{k \in \text{sequence}} Q_{av,k}}{K} \quad (22)$$

where, k indicates a k^{th} multimedia segment with the worst quality in the multimedia sequence, and $Q_{av,k}$ is multimedia quality of the k^{th} multimedia segment with the worst quality in the multimedia sequence.

[0093] Optionally, if weighted averaging is performed based on the weight value related to the multimedia quality of each multimedia segment of the K multimedia segments, the quality of the multimedia sequence may be determined by using the following equation (23):

$$Q = \frac{\sum_{k \in \text{sequence}} Q_{av,k} \cdot W_k}{\sum_{k \in \text{sequence}} W_k} \quad (23)$$

where, W_k is the weight value related to the multimedia quality.

[0094] Optionally, if weighted averaging is performed based on the weight value related to the time of each multimedia segment of the K multimedia segments, the quality of the multimedia sequence may be determined by using the following equation (24):

$$Q = \frac{\sum_{k \in \text{sequence}} Q_{av,k} \cdot W_{t_k}}{\sum_{k \in \text{sequence}} W_{t_k}} \quad (24)$$

where, W_{t_k} is the weight value related to the time.

[0095] Optionally, if weighted averaging is performed based on the weight value related to the multimedia quality and the time of each multimedia segment of the K multimedia segments, the quality of the multimedia sequence may be determined by using the following equation (25):

$$Q = \frac{\sum_{k \in \text{sequence}} Q_{av,k} \cdot W_k \cdot W_{t_k}}{\sum_{k \in \text{sequence}} W_k \cdot W_{t_k}} \quad (25)$$

[0096] Therefore, the method for evaluating multimedia quality according to this embodiment of the present invention, by dividing the multimedia sequence into segments, and then determining the multimedia quality of the multimedia sequence according to the multimedia quality of each multimedia segment, is convenient for determining the multimedia quality of the multimedia sequence according to the level of attention that is paid by the person to each multimedia segment, accords with the cognition characteristic of the person, and can increase the accuracy of evaluating the multimedia quality.

[0097] It should be understood that, in the embodiments of the present invention, a value of a sequence number of each step does not indicate an execution order, and the execution order of the steps should be determined according to a function and an inherent logic thereof, and should not constitute a limitation on an implementation process of the embodiments of the present invention.

[0098] With reference to FIG. 1 to FIG. 5, the foregoing describes in detail the method for evaluating multimedia quality according to the embodiments of the present invention. With reference to FIG. 6 to FIG. 9, the following describes an apparatus for evaluating multimedia quality according to the embodiments of the present invention.

[0099] FIG. 6 is a schematic block diagram of an apparatus 600 for evaluating multimedia quality according to an embodiment of the present invention. As shown in FIG. 6, the apparatus 600 includes:

a first obtaining module 610, configured to obtain reference quality of a video and final quality of the video of a multimedia sequence, and reference quality of an audio and final quality of the audio of the multimedia sequence;

a reference quality determining module 620, configured to determine reference quality of the multimedia sequence according to the reference quality of the video and the reference quality of the audio that are obtained by the first obtaining module 610;

a distortion value determining module 630, configured to determine a distortion value of the multimedia sequence according to the reference quality of the video, the final quality of the video, the reference quality of the audio, and the final quality of the audio that are obtained by the first obtaining module 610; and

a multimedia quality determining module 640, configured to determine multimedia quality of the multimedia sequence according to the reference quality of the multimedia sequence that is determined by the reference quality determining module 620 and the distortion value of the

multimedia sequence that is determined by the distortion value determining module 630.

[0100] The apparatus for evaluating multimedia quality according to this embodiment of the present invention, by determining the multimedia quality of the multimedia sequence according to the reference quality of the multimedia sequence and the distortion value of the multimedia sequence, can directly reflect the distortion of the multimedia sequence, accords with a subjective feeling of a person, and therefore can accurately and effectively evaluate the multimedia quality.

[0101] In this embodiment of the present invention, as shown in FIG. 7, optionally, the distortion value determining module 630 includes:

a first determining unit 631, configured to determine a distortion value of the video according to the reference quality of the video and the final quality of the video;

a second determining unit 632, configured to determine a distortion value of the audio according to the reference quality of the audio and the final quality of the audio; and

a third determining unit 633, configured to determine the distortion value of the multimedia sequence according to the distortion value of the video and the distortion value of the audio.

[0102] In this embodiment of the present invention, optionally, the third determining unit 633 includes:

a first determining subunit, configured to determine a distortion factor of the video according to the reference quality of the video and the distortion value of the video;

a second determining subunit, configured to determine a distortion factor of the audio according to the reference quality of the audio and the distortion value of the audio;

a third determining subunit, configured to determine a distortion factor of the multimedia sequence according to the distortion factor of the video and the distortion factor of the audio; and

a fourth determining subunit, configured to determine the distortion value of the multimedia sequence according to the reference quality of the multimedia sequence and the distortion factor of the multimedia sequence.

[0103] In this embodiment of the present invention, the third determining subunit is further configured to determine a distortion factor d_{av} of the multimedia sequence according to the following equation:

$$d_{av} = \frac{a_5 \cdot d_v + a_6 \cdot d_a}{1 + a_5 \cdot d_v + a_6 \cdot d_a} \quad \text{or}$$

$$d_{av} = a_5 + a_6 \cdot d_v + a_7 \cdot d_a,$$

where, d_v and d_a are respectively the distortion factor of the video and the distortion

factor of the audio, and a_5 , a_6 , and a_7 are constants.

[0104] In this embodiment of the present invention, as shown in FIG. 8, the apparatus 600 for evaluating multimedia quality further includes:

a second obtaining module 650, configured to obtain an effect factor of audio and video asynchronization of the multimedia sequence.

[0105] The multimedia quality determining module 640 is further configured to determine the multimedia quality of the multimedia sequence according to the reference quality of the multimedia sequence, the distortion value of the multimedia sequence, and the effect factor of audio and video asynchronization of the multimedia sequence that is obtained by the second obtaining module 650.

[0106] In this embodiment of the present invention, the multimedia quality determining module 640 is further configured to determine multimedia quality Q_{av} of the multimedia sequence according to the following equation:

$$Q_{av} = Q_{av}' - D_{av} \cdot f_5(\Delta T_{syn}) \quad \text{or}$$

$$Q_{av} = (Q_{av}' - D_{av}) \cdot f_6(\Delta T_{syn}),$$

where, Q_{av}' is the reference quality of the multimedia sequence, D_{av} is the distortion value of the multimedia sequence, $f_5(\Delta T_{syn})$ and $f_6(\Delta T_{syn})$ are effect factors of audio and video asynchronization of the multimedia sequence, and ΔT_{syn} is a time difference of audio and video asynchronization of the multimedia sequence; and a larger $|\Delta T_{syn}|$ indicates a larger $f_5(\Delta T_{syn})$ and a smaller $f_6(\Delta T_{syn})$.

[0107] The apparatus for evaluating multimedia quality according to this embodiment of the present invention, by determining the multimedia quality of the multimedia sequence according to the reference quality of the multimedia sequence and the distortion value of the multimedia sequence, can directly reflect the distortion of the multimedia sequence, accords with the subjective feeling of the person, and therefore can accurately and effectively evaluate the multimedia quality.

[0108] FIG. 9 is a schematic block diagram of an apparatus 900 for evaluating multimedia quality according to an embodiment of the present invention. As shown in FIG. 9, the apparatus 900 includes:

a segmenting module 910, configured to divide a multimedia sequence into N multimedia segments, where N is a positive integer and N is greater than or equal to 2;

an evaluating module 920, configured to evaluate multimedia quality of each multimedia segment of the N multimedia segments; and

a processing module 930, configured to determine multimedia quality of the multimedia

sequence according to the multimedia quality of each multimedia segment of the N multimedia segments.

[0109] The apparatus for evaluating multimedia quality according to this embodiment of the present invention, by dividing the multimedia sequence into segments, and determining the multimedia quality of the multimedia sequence according to the multimedia quality of each multimedia segment, is convenient for determining the multimedia quality of the multimedia sequence according to a level of attention that is paid by a person to each multimedia segment, accords with a cognition characteristic of the person, and can increase an accuracy of evaluating the multimedia quality.

[0110] In this embodiment of the present invention, optionally, the segmenting module 910 includes:

a first segmenting unit, configured to divide the multimedia sequence into the N multimedia segments according to a duration.

[0111] In this embodiment of the present invention, optionally, the segmenting module 910 includes:

a second segmenting unit, configured to divide the multimedia sequence into the N multimedia segments according to a level of multimedia quality, video quality, or audio quality.

[0112] In this embodiment of the present invention, the evaluating module 920 includes:

a first obtaining unit, configured to obtain reference quality of a video and final quality of the video of each multimedia segment, and reference quality of an audio and final quality of the audio of each multimedia segment;

a reference quality determining unit, configured to determine reference quality of each multimedia segment according to the reference quality of the video and the reference quality of the audio;

a distortion value determining unit, configured to determine a distortion value of each multimedia segment according to the reference quality of the video, the final quality of the video, the reference quality of the audio, and the final quality of the audio; and

an evaluating unit, configured to determine the multimedia quality of each multimedia segment according to the reference quality of each multimedia segment and the distortion value of each multimedia segment.

[0113] In this embodiment of the present invention, optionally, the evaluating module 920 further includes:

a second obtaining unit, configured to obtain an effect factor of audio and video asynchronization of each multimedia segment.

[0114] The evaluating unit is further configured to determine the multimedia quality of each multimedia segment according to the reference quality of each multimedia segment, the distortion value of each multimedia segment, and the effect factor of audio and video asynchronization of each multimedia segment.

5 **[0115]** In this embodiment of the present invention, the processing module 930 is further configured to determine the multimedia quality of the multimedia sequence by performing weighted averaging on multimedia quality of N or K multimedia segments, where the K multimedia segments are the K multimedia segments with worst multimedia quality of the N multimedia segments, and K is a positive integer and K is greater than or equal to 1 and less than N.

10 **[0116]** In this embodiment of the present invention, optionally, the processing module 930 includes:

a first processing unit, configured to determine the multimedia quality of the multimedia sequence based on an equal weight value, and by performing weighted averaging on the multimedia quality of the N or K multimedia segments.

15 **[0117]** In this embodiment of the present invention, optionally, the processing module 930 includes:

a second processing unit, configured to determine the multimedia quality of the multimedia sequence based on a weight value related to the multimedia quality of each multimedia segment of the N or K multimedia segments, and by performing weighted averaging on the multimedia quality of the N or K multimedia segments.

20 **[0118]** In this embodiment of the present invention, optionally, the processing module 930 includes:

a third processing unit, configured to determine the multimedia quality of the multimedia sequence based on a weight value related to a time of each multimedia segment of the N or K multimedia segments, and by performing weighted averaging on the multimedia quality of the N or K multimedia segments.

25 **[0119]** In this embodiment of the present invention, optionally, the processing module 930 includes:

a fourth processing unit, configured to determine the multimedia quality of the multimedia sequence based on a weight value related to the multimedia quality and a time of each multimedia segment of the N or K multimedia segments, and by performing weighted averaging on the multimedia quality of the N or K multimedia segments.

30 **[0120]** The apparatus for evaluating multimedia quality according to this embodiment of the present invention, by dividing the multimedia sequence into segments, and determining the

multimedia quality of the multimedia sequence according to the multimedia quality of each multimedia segment, is convenient for determining the multimedia quality of the multimedia sequence according to the level of attention that is paid by the person to each multimedia segment, accords with the cognition characteristic of the person, and can increase the accuracy of evaluating the multimedia quality.

[0121] It should be understood that, the term "and/or" in this embodiment of the present invention describes only an association relationship for describing associated objects and represents that three relationships may exist. For example, A and/or B may represent the following three cases: Only A exists, both A and B exist, and only B exists. In addition, the character "/" in this specification generally indicates an "or" relationship between the associated objects.

[0122] A person of ordinary skill in the art may be aware that, in combination with the examples described in the embodiments disclosed in this specification, units and algorithm steps may be implemented by electronic hardware or a combination of computer software and electronic hardware. Whether the functions are performed by hardware or software depends on particular applications and design constraint conditions of the technical solutions. A person skilled in the art may use different methods to implement the described functions for each particular application, but it should not be considered that the implementation goes beyond the scope of the present invention.

[0123] It may be clearly understood by a person skilled in the art that, for the purpose of convenient and brief description, for a detailed working process of the foregoing system, apparatus, and unit, reference may be made to a corresponding process in the foregoing method embodiments, and details are not described herein again.

[0124] In the several embodiments provided in the present application, it should be understood that the disclosed system, apparatus, and method may be implemented in other manners. For example, the described apparatus embodiment is merely exemplary. For example, the unit division is merely logical function division and may be other division in actual implementation. For example, a plurality of units or components may be combined or integrated into another system, or some features may be ignored or not performed. In addition, the displayed or discussed mutual couplings or direct couplings or communication connections may be implemented through some interfaces. The indirect couplings or communication connections between the apparatuses or units may be implemented in electronic, mechanical, or other forms.

[0125] The units described as separate parts may or may not be physically separate, and parts displayed as units may or may not be physical units, may be located in one position, or may be distributed on a plurality of network units. A part or all of the units may be selected as required to achieve the objectives of the solutions of the embodiments.

[0126] In addition, functional units in the embodiments of the present invention may be integrated into one processing unit, or each of the units may exist alone physically, or two or more units are integrated into one unit.

5 [0127] When the functions are implemented in a form of a software functional unit and sold or used as an independent product, the functions may be stored in a computer-readable storage medium. Based on such an understanding, the technical solutions of the present invention essentially, or the part contributing to the prior art, or a part of the technical solutions may be implemented in a form of a software product. The software product is stored in a storage medium, and includes several instructions for instructing a computer device (which may be a personal
10 computer, a server, or a network device) to perform all or a part of the steps of the methods described in the embodiments of the present invention. The foregoing storage medium includes: any medium that can store program code, such as a USB flash drive, a removable hard disk, a read-only memory (ROM, Read-Only Memory), a random access memory (RAM, Random Access Memory), a magnetic disk, or an optical disc.

15 [0128] The foregoing descriptions are merely specific implementation manners of the present invention, but are not intended to limit the protection scope of the present invention. Any variation or replacement readily figured out by persons skilled in the art within the technical scope disclosed in the present invention shall fall within the protection scope of the present invention. Therefore, the protection scope of the present invention shall be subject to the protection scope of the claims.

CLAIMS

What is claimed is:

1. A method for evaluating multimedia quality, comprising:

obtaining reference quality of a video and final quality of the video of a multimedia sequence,
5 and reference quality of an audio and final quality of the audio of the multimedia sequence;

determining reference quality of the multimedia sequence according to the reference quality of
the video and the reference quality of the audio;

determining a distortion value of the multimedia sequence according to the reference quality of
the video, the final quality of the video, the reference quality of the audio, and the final quality of
10 the audio; and

determining multimedia quality of the multimedia sequence according to the reference quality
of the multimedia sequence and the distortion value of the multimedia sequence.

2. The method according to claim 1, wherein the determining a distortion value of the
multimedia sequence according to the reference quality of the video, the final quality of the video,
15 the reference quality of the audio, and the final quality of the audio comprises:

determining a distortion value of the video according to the reference quality of the video and
the final quality of the video;

determining a distortion value of the audio according to the reference quality of the audio and
the final quality of the audio; and

20 determining the distortion value of the multimedia sequence according to the distortion value
of the video and the distortion value of the audio.

3. The method according to claim 2, wherein the determining the distortion value of the
multimedia sequence according to the distortion value of the video and the distortion value of the
audio comprises:

25 determining a distortion factor of the video according to the reference quality of the video and
the distortion value of the video;

determining a distortion factor of the audio according to the reference quality of the audio and
the distortion value of the audio;

determining a distortion factor of the multimedia sequence according to the distortion factor of
30 the video and the distortion factor of the audio; and

determining the distortion value of the multimedia sequence according to the reference quality
of the multimedia sequence and the distortion factor of the multimedia sequence.

4. The method according to claim 3, wherein the determining a distortion factor of the

multimedia sequence according to the distortion factor of the video and the distortion factor of the audio comprises:

determining a distortion factor d_{av} of the multimedia sequence according to the following equation:

$$d_{av} = \frac{a_5 \cdot d_v + a_6 \cdot d_a}{1 + a_5 \cdot d_v + a_6 \cdot d_a} \text{ or}$$

$$d_{av} = a_5 + a_6 \cdot d_v + a_7 \cdot d_a,$$

wherein, d_v and d_a are respectively the distortion factor of the video and the distortion factor of the audio, and a_5 , a_6 , and a_7 are constants.

5. A method for evaluating multimedia quality, comprising:

dividing a multimedia sequence into N multimedia segments, wherein N is a positive integer and N is greater than or equal to 2;

evaluating multimedia quality of each multimedia segment of the N multimedia segments; and

determining multimedia quality of the multimedia sequence according to the multimedia quality of each multimedia segment of the N multimedia segments.

6. The method according to claim 5, wherein the evaluating multimedia quality of each multimedia segment of the N multimedia segments comprises:

obtaining reference quality of a video and final quality of the video of each multimedia segment, and reference quality of an audio and final quality of the audio of each multimedia segment;

determining reference quality of each multimedia segment according to the reference quality of the video and the reference quality of the audio;

determining a distortion value of each multimedia segment according to the reference quality of the video, the final quality of the video, the reference quality of the audio, and the final quality of the audio; and

determining the multimedia quality of each multimedia segment according to the reference quality of each multimedia segment and the distortion value of each multimedia segment.

7. The method according to claim 6, wherein the evaluating multimedia quality of each multimedia segment of the N multimedia segments further comprises:

obtaining an effect factor of audio and video asynchronization of each multimedia segment;

and

determining the multimedia quality of each multimedia segment according to the reference quality of each multimedia segment and the distortion value of each multimedia segment comprises:

determining the multimedia quality of each multimedia segment according to the reference

quality of each multimedia segment, the distortion value of each multimedia segment, and the effect factor of audio and video asynchronization of each multimedia segment.

8. The method according to any one of claims 5 to 7, wherein the dividing a multimedia sequence into N multimedia segments comprises:

- 5 dividing the multimedia sequence into the N multimedia segments according to a duration; or
dividing the multimedia sequence into the N multimedia segments according to a level of multimedia quality, video quality, or audio quality.

9. The method according to any one of claims 5 to 7, wherein the determining multimedia quality of the multimedia sequence according to the multimedia quality of each multimedia segment
10 of the N multimedia segments comprises:

determining the multimedia quality of the multimedia sequence by performing weighted averaging on multimedia quality of N or K multimedia segments, wherein the K multimedia segments are the K multimedia segments with worst multimedia quality of the N multimedia segments, and K is a positive integer and K is greater than or equal to 1 and less than N.

15 10. An apparatus for evaluating multimedia quality, comprising:

a first obtaining module, configured to obtain reference quality of a video and final quality of the video of a multimedia sequence, and reference quality of an audio and final quality of the audio of the multimedia sequence;

20 a reference quality determining module, configured to determine reference quality of the multimedia sequence according to the reference quality of the video and the reference quality of the audio that are obtained by the first obtaining module;

25 a distortion value determining module, configured to determine a distortion value of the multimedia sequence according to the reference quality of the video, the final quality of the video, the reference quality of the audio, and the final quality of the audio that are obtained by the first obtaining module; and

a multimedia quality determining module, configured to determine multimedia quality of the multimedia sequence according to the reference quality of the multimedia sequence that is determined by the reference quality determining module and the distortion value of the multimedia sequence that is determined by the distortion value determining module.

30 11. The apparatus according to claim 10, wherein the distortion value determining module comprises:

a first determining unit, configured to determine a distortion value of the video according to the reference quality of the video and the final quality of the video;

a second determining unit, configured to determine a distortion value of the audio according to

the reference quality of the audio and the final quality of the audio; and

a third determining unit, configured to determine the distortion value of the multimedia sequence according to the distortion value of the video and the distortion value of the audio.

12. The apparatus according to claim 11, wherein the third determining unit comprises:

5 a first determining subunit, configured to determine a distortion factor of the video according to the reference quality of the video and the distortion value of the video;

a second determining subunit, configured to determine a distortion factor of the audio according to the reference quality of the audio and the distortion value of the audio;

10 a third determining subunit, configured to determine a distortion factor of the multimedia sequence according to the distortion factor of the video and the distortion factor of the audio; and

a fourth determining subunit, configured to determine the distortion value of the multimedia sequence according to the reference quality of the multimedia sequence and the distortion factor of the multimedia sequence.

13. The apparatus according to claim 12, wherein the third determining subunit is further
15 configured to determine a distortion factor d_{av} of the multimedia sequence according to the following equation:

$$d_{av} = \frac{a_5 \cdot d_v + a_6 \cdot d_a}{1 + a_5 \cdot d_v + a_6 \cdot d_a} \text{ or}$$

$$d_{av} = a_5 + a_6 \cdot d_v + a_7 \cdot d_a,$$

20 wherein, d_v and d_a are respectively the distortion factor of the video and the distortion factor of the audio, and a_5 , a_6 , and a_7 are constants.

14. An apparatus for evaluating multimedia quality, comprising:

a segmenting module, configured to divide a multimedia sequence into N multimedia segments, wherein N is a positive integer and N is greater than or equal to 2;

25 an evaluating module, configured to evaluate multimedia quality of each multimedia segment of the N multimedia segments; and

a processing module, configured to determine multimedia quality of the multimedia sequence according to the multimedia quality of each multimedia segment of the N multimedia segments.

15. The apparatus according to claim 14, wherein the evaluating module comprises:

30 a first obtaining unit, configured to obtain reference quality of a video and final quality of the video of each multimedia segment, and reference quality of an audio and final quality of the audio of each multimedia segment;

a reference quality determining unit, configured to determine reference quality of each multimedia segment according to the reference quality of the video and the reference quality of the

audio;

a distortion value determining unit, configured to determine a distortion value of each multimedia segment according to the reference quality of the video, the final quality of the video, the reference quality of the audio, and the final quality of the audio; and

5 an evaluating unit, configured to determine the multimedia quality of each multimedia segment according to the reference quality of each multimedia segment and the distortion value of each multimedia segment.

16. The apparatus according to claim 15, wherein the evaluating module further comprises:

10 a second obtaining unit, configured to obtain an effect factor of audio and video asynchronization of each multimedia segment; and

the evaluating unit is further configured to determine the multimedia quality of each multimedia segment according to the reference quality of each multimedia segment, the distortion value of each multimedia segment, and the effect factor of audio and video asynchronization of each multimedia segment.

15 17. The apparatus according to any one of claims 14 to 16, wherein the segmenting module comprises:

a first segmenting unit, configured to divide the multimedia sequence into the N multimedia segments according to a duration; or

20 a second segmenting unit, configured to divide the multimedia sequence into the N multimedia segments according to a level of multimedia quality, video quality, or audio quality.