

METHOD AND DEVICE FOR DETERMINING INTER-CUT TIME BUCKET IN AUDIO OR VIDEO

TECHNICAL FIELD

[0001] The present application relates to the field of data processing technologies,
5 and in particular, to a method and a device for determining an inter-cut time bucket in
an audio or a video.

BACKGROUND

[0002] Television program playback is an important component of video content
provided on many video websites. A recorded television program video usually
10 includes a program period and an inter-cut time bucket (e.g. a commercial break). The
program period is a period for playing program content. The inter-cut time bucket is a
period for playing non-program content, for example, an advertisement period in a
movie or an episode of TV series, or an advertisement period or a live broadcast room
comment period for filling a half-time break in a live program of a sports game.

15 [0003] When watching a playback video, few users wait patiently during the
inter-cut time bucket. Instead, the users usually skip the inter-cut time bucket by
dragging a progress bar. The users usually need to adjust the progress bar back and
forth because it is difficult to accurately skip the inter-cut time bucket through
one-time dragging, which affects user experience. Therefore, the inter-cut time bucket
20 is usually removed from the program video when the playback video is produced.

[0004] In the existing technology, a video producer manually identifies an
inter-cut time bucket in a video, and then cuts the inter-cut time bucket off the video.
The present method is tedious, time-consuming, and inefficient, and therefore, is hard
to satisfy the demand for putting the playback video online in time.

25

SUMMARY

[0005] In view of this, the present application provides a method for determining
an inter-cut time bucket in an audio or a video, including the following: determining,

as a candidate segmentation point, at least one of a time point at which feature media data appears in an audio or a video or a time point at which feature media data disappears from an audio or a video; and determining a candidate segmentation point satisfying a predetermined start point condition as a start point, determining a
 5 candidate segmentation point satisfying a predetermined end point condition as an end point, and determining a period between the start point and the end point as an inter-cut time bucket.

[0006] The present application further provides a device for determining an inter-cut time bucket in an audio or a video, including: a candidate segmentation point
 10 unit, configured to use, as a candidate segmentation point, at least one of a time point at which feature media data appears in an audio or a video or a time point at which feature media data disappears from an audio or a video; and a start point and end point unit, configured to use a candidate segmentation point satisfying a predetermined start point condition as a start point, use a candidate segmentation point satisfying a
 15 predetermined end point condition as an end point, and use a period between the start point and the end point as an inter-cut time bucket.

[0007] It can be seen from the previous technical solutions that, in the present implementations of the present application, whether feature media data is presented in an audio or a video at a certain time point is used as a basis for distinguishing between
 20 a program period and an inter-cut time bucket. In at least one of a time point at which the feature media data appears or a time point at which the feature media data disappears, one satisfying a predetermined start point condition is used as a start point of the inter-cut time bucket, and the other satisfying a predetermined end point condition is used as an end point of the inter-cut time bucket. The inter-cut time
 25 bucket in the audio/video can be automatically identified without manual identification. As such, during playback video production, manual workloads on the producer can be greatly alleviated, a period for producing the playback video can be shortened, and improving production efficiency.

BRIEF DESCRIPTION OF DRAWINGS

[0008] FIG. 1 is a flowchart illustrating a method for determining an inter-cut time bucket in an audio or a video, according to an implementation of the present application;

[0009] FIG. 2 is an example diagram illustrating a video picture with a feature image displayed, according to an application example of the present application;

[0010] FIG. 3 is an example diagram illustrating a playback progress bar in which an area corresponding to an inter-cut time bucket is presented by using a different color, according to an application example of the present application;

[0011] FIG. 4 is a structural diagram illustrating hardware of a device for running an implementation of the present application; and

[0012] FIG. 5 is a logic structural diagram illustrating a device for determining an inter-cut time bucket in an audio or a video, according to an implementation of the present application.

DESCRIPTION OF IMPLEMENTATIONS

[0013] In an audio or a video including an inter-cut time bucket, especially, a recorded video of a television program, generally, audio/video content in a program period or the inter-cut time bucket has obvious identifiable features. For example, in an advertisement insertion period in the middle of a movie or TV series, a countdown of the advertisement insertion period is constantly displayed, "To be continued later" is constantly displayed, etc. For another example, generally, a scoreboard is constantly displayed in a program period of a sports game. The at least one of appearance or disappearance of these identifiable features can be used as a basis for division between a program period and an inter-cut time bucket.

[0014] Therefore, an implementation of the present application provides a new method for determining an inter-cut time bucket in an audio or a video. The method includes the following: determining, as a candidate segmentation point between the program period and the inter-cut time bucket, at least one of a time point at which feature media data of the identifiable feature appears or a time point at which feature media data of the identifiable feature disappears; and determining a candidate segmentation point satisfying a predetermined start point condition as a start point of the inter-cut time bucket, and determining a candidate segmentation point satisfying a predetermined end point condition as an end point of the inter-cut time bucket. As such, the inter-cut time bucket in the audio/video can be automatically identified without manual identification. Manual workloads on the producer can be greatly alleviated, and a speed for producing the video can be fastened, improving production

efficiency, and alleviating a problem in the existing technology.

[0015] The present implementations of the present application can be applied to any device having computing and storage capabilities, such as a mobile phone, a tablet computer, a personal computer (PC), a notebook, or a server. Each function in the present implementations of the present application can be further implemented by a logical node running on two or more devices.

[0016] FIG. 1 is a flowchart illustrating a method for determining an inter-cut time bucket in an audio or a video, according to an implementation of the present application.

[0017] Step 110: Use, as a candidate segmentation point, at least one of a time point at which feature media data appears in an audio or a video or a time point at which feature media data disappears from an audio or a video.

[0018] In the present implementation of the present application, a program period and an inter-cut time bucket in the audio/video can be divided based on a demand in an actual application scenario. This is not limited. For example, a pause period in a sport game program can be used as a program period, or can be used as an inter-cut time bucket.

[0019] At least one of an image feature or a sound feature that appears or disappears when switching is performed between the program period and the inter-cut time bucket can be used as the feature media data after the program period and the inter-cut time bucket are determined. The at least one of the time point at which the feature media data appears in the audio/video or the time point at which the feature media data disappears from the audio/video is identified as the candidate segmentation point.

[0020] For the video, the feature media data can be a feature image in a video picture or a feature sound in a video sound. For the audio, the feature media data is generally a feature sound in the audio. It is worthwhile to note that the feature image can be all or a part of a video picture and the feature sound can be all or a part of a band of the audio, for example, a sound with at least one of a predetermined volume, tone, or timbre. This is not limited in the present implementation of the present application.

[0021] The video picture, the video sound, or the audio sound is analyzed to identify whether the current video picture, the video sound, or the audio sound has the feature media data, thereby further identifying the at least one of the time point at

which the feature media data appears or the time point at which the feature media data disappears.

[0022] When the feature image is a part of the video picture and the feature image always appears in a predetermined area in the picture, at least one of a color ratio or an image shape of the predetermined area in the video picture can be analyzed, and at least one of the time point at which the feature image appears in the predetermined area or the time point at which the feature image disappears from the predetermined area is identified. A scoreboard in a sports game video is used as an example. Assuming that the scoreboard is not displayed in a certain frame but is displayed in a next frame, a time point between the two frames is a candidate segmentation point.

[0023] There can be one or more pieces of feature media data for a certain audio/video or a certain type of audio/video. For example, for a sports game video, a scoreboard can be used as the feature media data, or both a scoreboard and a long whistling in a game pause can be used as the feature media data. For another example, for a movie or a television program video, at least one of displayed "To be continued later" or a displayed countdown for an advertisement period can be used as the feature media data.

[0024] Based on a method for presenting feature media data in an actual application scenario, at least one of a time point at which certain feature media data appears or a time point at which certain feature media data disappears can be used as a candidate segmentation point. For example, for a countdown displayed in an advertisement insertion period, if a countdown is displayed throughout an advertisement insertion period in a certain video, both a time point at which the countdown appears and a time point at which the countdown disappears can be used as candidate segmentation points. If a countdown is displayed only before an advertisement insertion period is about to end in a certain video, instead of a time point at which the displayed countdown appears, a time point at which the displayed countdown disappears is used as a candidate segmentation point.

[0025] Step 120: Use a candidate segmentation point satisfying a predetermined start point condition as a start point, use a candidate segmentation point satisfying a predetermined end point condition as an end point, and use a period between the start point and the end point as an inter-cut time bucket.

[0026] For the candidate segmentation point, based on a method for presenting selected feature media data in an audio or a video, a change representing a start of the

inter-cut time bucket is used as the predetermined start point condition, and a change representing an end of the inter-cut time bucket is used as the predetermined end point condition, to determine a start time point and an end time point of one or more spot periods from candidate segmentation points. As such, each spot period is determined from the audio/video.

[0027] For the video, the feature media data can be one or more feature images in a video picture. When a certain feature image can be used for determining a start point of an inter-cut time bucket, a predetermined start point condition can be that the feature image appears or disappears. When a certain feature image can be used for determining an end point of an inter-cut time bucket, a predetermined end point condition can be that the feature image appears or disappears. When a certain feature image can be used for determining both a start point and an end point of an inter-cut time bucket, a predetermined start point condition can be that the feature image appears or disappears, and a corresponding predetermined end point condition can be that the feature image disappears or appears.

[0028] For example, when "To be continued later" displayed in the advertisement period can generally be used for determining a start of the inter-cut time bucket, a corresponding predetermined start point condition is that "To be continued later" appears; when a countdown displayed in the advertisement period is used for determining an end of the inter-cut time bucket, a corresponding predetermined end point condition can be that the countdown disappears.

[0029] Similarly, the feature media data can be one or more feature sounds for the audio/video. When a certain feature sound can be used for determining a start point of an inter-cut time bucket, a predetermined start point condition can be that the feature sound appears or disappears. When a certain feature sound can be used for determining an end point of an inter-cut time bucket, a predetermined end point condition can be that the feature sound appears or disappears. When a certain feature sound can be used for determining both a start point and an end point of an inter-cut time bucket, a predetermined start point condition can be that the feature sound appears or disappears, and a corresponding predetermined end point condition can be that the feature sound disappears or appears.

[0030] When two or more pieces of feature media data are used, at least one of a relatively complex predetermined start point condition or a relatively complex predetermined end point condition can be used based on a method for displaying the

feature media data in an actual application scenario. For example, when a playback video of a sports game is produced, a pause period, an advertisement period, and a live broadcast room period of the game are used as spot periods, and two types of feature media data: a scoreboard and a long whistling for a pause can be used. In the present scenario, the scoreboard is not displayed in the advertisement period or the live broadcast room period but is displayed in the pause period. In addition, the advertisement period or the live broadcast room period is usually after the pause period starts, and the pause period may end or not end after the advertisement period or the live broadcast room period ends. As such, a predetermined start point condition can be that the long whistling for a pause appears and the scoreboard does not disappear in a previous first predetermined waiting period, or that the scoreboard disappears; a predetermined end point condition can be that the scoreboard appears and the long whistling for a pause does not appear in a following second predetermined waiting period, or that the long whistling for a pause appears and a period between a time point at which the long whistling for a pause appears and a time point at which the long whistling for a pause appears last time does not exceed a third predetermined waiting period.

[0031] The candidate segmentation point satisfying the predetermined start point condition is used as the start point of the inter-cut time bucket, the candidate segmentation point satisfying the predetermined end point condition is used as the end point of the inter-cut time bucket, and the period between the start point and the end point is the inter-cut time bucket. It is worthwhile to note that there can be one or more spot periods in the audio/video, and each spot period can be determined based on the method in the present implementation of the present application.

[0032] After the inter-cut time bucket in the audio/video is determined, different colors can be used in a progress bar to show an area corresponding to the inter-cut time bucket and an area corresponding to a non-spot period (namely, the program period) when the audio/video is played. As such, a user can conveniently jump to a next program period by dragging the progress bar. Alternatively, all spot periods may not be played and only the program period can be played when the audio/video is played. Alternatively, the inter-cut time bucket can be deleted from the audio/video, and a video version obtained after the inter-cut time bucket is deleted and only the program period is left is used during audio/video playing.

[0033] It can be seen that, in the present implementation of the present application,

an identifiable feature can be used as the feature media data for distinguishing between the program period and the inter-cut time bucket. In at least one of the time point at which the feature media data appears or the time point at which the feature media data disappears, one satisfying the predetermined start point condition is used as the start point of the inter-cut time bucket, and the other satisfying the predetermined end point condition is used as the end point of the inter-cut time bucket. The inter-cut time bucket in the audio/video can be automatically identified without manual identification. As such, manual workloads on the producer can be greatly alleviated, improving a video production speed and video production efficiency.

[0034] In an application example of the present application, after recording a television live broadcast NBA game video, the producer wants to make the video as a playback video. In addition to a game period, a television live broadcast program further includes an advertisement period and a live broadcast room period. During video production, a game period is a program period in the video, and an advertisement period and a live broadcast room period are spot periods. In the game period of the video, a scoreboard is always displayed at the lower right corner of a picture, as shown in FIG. 2, while the scoreboard is not displayed in the advertisement period and the live room period. Therefore, the producer determines a display area including the scoreboard at the lower right corner of the picture as a predetermined area, and determines the displayed scoreboard as a feature image.

[0035] The device running the present implementation of the present application extracts each frame from the video, and analyzes a color ratio and an image shape of a predetermined area in the frame. As such, whether the scoreboard is displayed in the frame can be identified. When the scoreboard is displayed in a previous frame but is not displayed in a following frame, or the scoreboard is not displayed in a previous frame but is displayed in a following frame, a time point between the two frames is used as a candidate segmentation point. A candidate segmentation point at which the scoreboard disappears (a predetermined start point condition) is used as a start point of an inter-cut time bucket, and a candidate segmentation point at which the scoreboard appears (a predetermined end point condition) is used as an end point of the inter-cut time bucket, so that spot periods are divided in the video.

[0036] The device running the present implementation of the present application can delete the inter-cut time bucket from the video, and use a video obtained after the deletion as a playback video to be played for the user. Alternatively, for example, as

shown in FIG. 3, the recorded video can be used as a playback video, and different colors are used in a progress bar to show an area corresponding to the inter-cut time bucket and an area corresponding to the program period when the video is played, so that the user can accurately skip the inter-cut time bucket.

5 [0037] Corresponding to the previous procedure implementation, an implementation of the present application further provides a device for determining an inter-cut time bucket in an audio or a video. The device can be implemented by software, can be implemented by hardware, or can be implemented by a combination of hardware and software. Software implementation is used as an example. As a
10 logical apparatus, the software is formed by reading a corresponding computer program instruction to a memory by using a central processing unit (CPU) in a device in which the software is located and running the computer program instruction. In terms of hardware implementation, in addition to the CPU, the memory, and the nonvolatile memory shown in FIG. 4, the device in which the device for determining
15 an inter-cut time bucket in an audio or a video is located generally further includes at least one of the following: other hardware such as a chip for sending and receiving radio signals, or other hardware such as a card configured to implement a network communication function.

[0038] FIG. 5 shows a device for determining an inter-cut time bucket in an audio
20 or a video according to an implementation of the present application. The device includes a candidate segmentation point unit and a start point and end point unit. The candidate segmentation point unit is configured to use, as a candidate segmentation point, at least one of a time point at which feature media data appears in an audio or a video or a time point at which feature media data disappears from an audio or a video.
25 The start point and end point unit is configured to use a candidate segmentation point satisfying a predetermined start point condition as a start point, use a candidate segmentation point satisfying a predetermined end point condition as an end point, and use a period between the start point and the end point as an inter-cut time bucket.

[0039] Optionally, the feature media data includes a feature image in a video
30 picture, and the predetermined start point condition includes that the feature image appears in or disappears from the video picture.

[0040] Optionally, the feature media data includes a feature image in a video picture, and the predetermined end point condition includes that the feature image disappears from or appears in the video picture.

[0041] Optionally, the feature image is located in a predetermined area of the video picture, and the candidate segmentation point unit is configured to analyze at least one of a color ratio or an image shape of the predetermined area in the video picture, and identify, as the candidate segmentation point, at least one of the time
 5 point at which the feature image appears in the predetermined area or the time point at which the feature image disappears from the predetermined area.

[0042] Optionally, the feature image includes a scoreboard in a sports game video.

[0043] Optionally, the feature media data includes a feature sound in the audio/video, and the predetermined start point condition includes a time point at
 10 which the feature sound appears in or disappears from the audio/video.

[0044] Optionally, the feature media data includes a feature sound in the audio/video, and the predetermined end point condition includes a time point at which the feature sound disappears from or appears in the audio/video.

[0045] Optionally, the feature sound includes a sound with at least one of a
 15 predetermined volume, tone, or timbre.

[0046] Optionally, the device further includes: a unit for playing and changing an audio or a video, configured to use different colors in a progress bar to show an area corresponding to the inter-cut time bucket and an area corresponding to a non-spot period when the audio/video is played; or skip displaying the inter-cut time bucket
 20 when the audio/video is played; or delete the inter-cut time bucket from the audio/video.

[0047] The previous descriptions are merely examples of the present implementations of the present application, but are not intended to limit the present application. Any modification, equivalent replacement, or improvement made without
 25 departing from the spirit and principle of the present application should fall within the protection scope of the present application.

[0048] In a typical configuration, a computing device includes one or more central processing units (CPU), one or more input/output interfaces, one or more network interfaces, and one or more memories.

[0049] The memory can include a non-persistent memory, a random access memory (RAM), a nonvolatile memory, etc. in a computer readable medium, for example, a read-only memory (ROM) or a flash memory (flash RAM). The memory
 30 is an example of the computer readable medium.

[0050] The computer readable medium includes persistent, non-persistent,

movable, and unmovable media that can store information by using any method or technology. The information can be a computer readable instruction, a data structure, a program module, or other data. Examples of the computer storage medium include but are not limited to: a parameter random access memory (PRAM), a static random
5 access memory (SRAM), a dynamic random access memory (DRAM) or another type of random access memory (RAM), a read-only memory (ROM), an electrically erasable programmable read-only memory (EEPROM), a flash memory or another memory technology, a compact disc read-only memory (CD-ROM), a digital versatile disc (DVD) or another optical storage, a magnetic cassette, a magnetic tape, a
10 magnetic tape/magnetic disk memory or another magnetic storage device, or any other non-transmission medium that can be used to store information that can be accessed by a computing device. As described in the present application, the computer readable medium does not include computer readable transitory media such as a modulated data signal and a carrier.

15 **[0051]** It is worthwhile to further note that, the terms "include", "comprise", or their any other variants are intended to cover a non-exclusive inclusion, so a process, a method, a product or a device that includes a list of elements not only includes those elements but also includes other elements which are not expressly listed, or further includes elements inherent to such process, method, product or device. Without more
20 constraints, an element preceded by "includes a ..." does not preclude the existence of additional identical elements in the process, method, product or device that includes the element.

[0052] A person skilled in the art should understand that an implementation of the present application can be provided as a method, a system, or a computer program
25 product. Therefore, the present application can use a form of hardware only implementations, software only implementations, or implementations with a combination of software and hardware. In addition, the present application can use a form of a computer program product that is implemented on one or more computer-usable storage media (including but not limited to a magnetic disk memory,
30 a CD-ROM, an optical memory, etc.) that include computer-usable program code.

CLAIMS

What is claimed is:

1. A method for determining an inter-cut time bucket in an audio or a video, comprising:

5 determining, as a candidate segmentation point, at least one of a time point at which feature media data appears in an audio or a video, or a time point at which feature media data disappears from an audio or a video; and

 determining a candidate segmentation point satisfying a predetermined start point condition as a start point, determining a candidate segmentation point satisfying
10 a predetermined end point condition as an end point, and determining a period between the start point and the end point as an inter-cut time bucket.

2. The method according to claim 1, wherein the feature media data comprises a feature image in a video picture, and wherein the predetermined start point condition
15 comprises that the feature image appears in or disappears from the video picture.

3. The method according to claim 1, wherein the feature media data comprises a feature image in a video picture, and wherein the predetermined end point condition comprises that the feature image disappears from or appears in the video picture.
20

4. The method according to claim 2 or 3, wherein the feature image is located in a predetermined area in the video picture; and

 wherein the determining, as a candidate segmentation point, at least one of a time point at which feature media data appears in an audio or a video, or a time point at
25 which feature media data disappears from an audio or a video comprises:

 analyzing at least one of a color ratio or an image shape of the predetermined area in the video picture; and

 identifying, as the candidate segmentation point, at least one of the time point at which the feature image appears in the predetermined area or the time point at
30 which the feature image disappears from the predetermined area.

5. The method according to claim 1, wherein the feature image comprises a

scoreboard in a sports game video.

6. The method according to claim 1, wherein the feature media data comprises a feature sound in the video or the audio, and wherein the predetermined start point
5 condition comprises a time point at which the feature sound appears in or disappears from the video or the audio.

7. The method according to claim 1, wherein the feature media data comprises a feature sound in the video or the audio, and wherein the predetermined end point
10 condition comprises a time point at which the feature sound disappears from or appears in the video or the audio.

8. The method according to claim 6 or 7, wherein the feature sound comprises a sound with at least one of a predetermined volume, tone, or timbre.
15

9. The method according to claim 1, wherein the method further comprises:
using different colors in a progress bar to show an area corresponding to the inter-cut time bucket and an area corresponding to a non-spot period when the audio or video is played; or
20 skipping playing the inter-cut time bucket when the audio or video is played; or deleting the inter-cut time bucket from the audio or video.

10. A device for determining an inter-cut time bucket in an audio or a video, comprising:
25 a candidate segmentation point unit, configured to determine, as a candidate segmentation point, at least one of a time point at which feature media data appears in an audio or a video, or a time point at which feature media data disappears from an audio or a video; and
a start point and end point unit, configured to:
30 determine a candidate segmentation point satisfying a predetermined start point condition as a start point,
determine a candidate segmentation point satisfying a predetermined end point condition as an end point, and
determine a period between the start point and the end point as an inter-cut

time bucket.

11. The device according to claim 10, wherein the feature media data comprises a feature image in a video picture, and wherein the predetermined start point condition
5 comprises that the feature image appears in or disappears from the video picture.

12. The device according to claim 10, wherein the feature media data comprises a feature image in a video picture, and wherein the predetermined end point condition comprises that the feature image disappears from or appears in the video picture.
10

13. The device according to claim 11 or 12, wherein the feature image is located in a predetermined area in the video picture; and
wherein the candidate segmentation point unit is configured to:
analyze at least one of a color ratio or an image shape of the predetermined
15 area in the video picture; and
identify, as the candidate segmentation point, at least one of the time point at which the feature image appears in the predetermined area or the time point at which the feature image disappears from the predetermined area.

20 14. The device according to claim 10, wherein the feature image comprises a scoreboard in a sports game video.

15. The device according to claim 10, wherein the feature media data comprises a feature sound in the video or the audio, and wherein the predetermined start point
25 condition comprises a time point at which the feature sound appears in or disappears from the video or the audio.

16. The device according to claim 10, wherein the feature media data comprises a feature sound in the video or the audio, and wherein the predetermined end point
30 condition comprises a time point at which the feature sound disappears from or appears in the video or the audio.

17. The device according to claim 15 or 16, wherein the feature sound comprises a sound with at least one of a predetermined volume, tone, or timbre.

18. The device according to claim 10, wherein the device further comprises: a unit for playing and changing an audio or a video, configured to use different colors in a progress bar to show an area corresponding to the inter-cut time bucket and an area
5 corresponding to a non-spot period when the audio or video is played; or skip displaying the inter-cut time bucket when the audio or video is played; or delete the inter-cut time bucket from the audio or video.