WHAT IS CLAIMED IS

1. A segmentation method, comprising:

receiving a video data record, wherein the video data record comprises an image data record and an audio data record;

performing segmentation processing for the image data record to generate at least one image paragraph mark;

performing segmentation processing for the audio data record to generate at least one audio paragraph mark; and

calculating a time difference value between an image marking time of the at least one image paragraph mark and an audio marking time of the at least one audio paragraph mark and generating at least one video content mark,

wherein the time difference value is compared with a threshold to determine whether or not to use the image marking time and the audio marking time for generating the at least one video content mark, and

wherein when the time difference value is less than the threshold, the at least one video content mark is generated by the time average of the image marking time and the audio marking time.

2. The segmentation method of claim 1, wherein performing segmentation processing for the image data record to generate the at least one image paragraph mark, further comprising:

selecting M unit basis of the image data record, and grouping the selected image data record into a first image paragraph;

determining content of the first image paragraph to generate an image content result, wherein the image content result comprises a dynamic content and a static content; and

detecting a change content for the image data record based on the image content result and generating the at least one image paragraph mark according to time position of the change content.

3. The segmentation method of claim 2, wherein determining content of the first image paragraph to generate the image content result, further comprising:

selecting T unit basis of the first image paragraph, and calculating similarity of the image data record in T unit basis to generate an image difference result;

if the image difference result is larger than a first image threshold, determining content the first image paragraph as the dynamic content; and

if the image difference result is less than the first image threshold, determining content the first image paragraph as the static content.

4. The segmentation method of claim 2, wherein detecting the change content for the image data record based on the image content result and generating the at least one image paragraph mark according to time position of the change content, further comprising:

if the first image paragraph is the dynamic content, calculating similarity between an Mth unit basis of the image data record and an (M+1)th unit basis of the image data record to generate an image difference value;

if the image difference value is larger than a second image threshold, merging the (M+1)th unit basis of the image data record into the first image paragraph; and

if the image difference value is less than the second image threshold, generating the at least one image paragraph mark in time position of the (M+1)th unit basis of the image data record, and selecting M unit basis of the

image data record, and grouping the selected image data record into a second image paragraph.

5. The segmentation method of claim 2, wherein detecting the change content for the image data record based on the image content result and generating the at least one image paragraph mark according to time position of the change content, further comprising:

if the first image paragraph is the static content, calculating similarity between an Mth unit basis of the image data record and an (M+1)th unit basis of the image data record to generate an image difference value;

if the image difference value is less than a second image threshold, merging the (M+1)th unit basis of the image data record into the first image paragraph; and

if the image difference value is larger than the second image threshold, generating the at least one image paragraph mark in time position of the (M+1)th unit basis of the image data record, and selecting M unit basis of the image data record, and grouping the selected image data record into a second image paragraph.

6. The segmentation method of claim 1, wherein performing segmentation processing for the audio data record to generate the at least one audio paragraph mark, further comprising:

transforming the audio data record into an audio time domain signal and an audio frequency domain signal, respectively;

selecting a time domain section from the audio time domain signal, and determining whether amplitude of the time domain section is less than a first threshold, if amplitude of the time domain section is less than the first threshold, generating the at least one audio paragraph mark; and

selecting a first frequency domain section and a second frequency domain section from the audio frequency domain signal, and determining whether a difference value of spectral intensity between the first frequency domain section and the second frequency domain section is greater than a second threshold, if the difference value of spectral intensity between the first frequency domain section and the second frequency domain section is larger than the second threshold, generating the at least one audio paragraph mark.

7. A segmentation system, comprising:

a storage unit, configured to storge a video data record and at least one video content mark; and

a processor, electrically connected to the storage unit, and configured to receive the video data record, wherein the video data record comprises an image data record and an audio data record, the processor further comprises:

an image segmentation unit, configured to perform segmentation processing for the image data record to generate at least one image paragraph mark;

an audio segmentation unit, electrically connected to the image segmentation unit, and configured to perform segmentation processing for the audio data record to generate at least one audio paragraph mark; and

a paragraph mark generating unit, electrically connected to the image segmentation unit and the audio segmentation unit, and configured to calculate a time difference value between an image marking time of the at least one image paragraph mark and an audio marking time of the at least one audio paragraph mark and to generate the at least one video content mark,

wherein the time difference value is compared with a threshold to deternine whether or not to use the image marking time and the audio marking time for generating the at least one video content mark, and

wherein when the time difference value is less than the threshold, the at least one video content mark is generated by the time average of the image marking time and the audio marking time.

8. The segmentation system of claim 7, wherein wherein the image segmentation unit is further configured for executing following steps:

selecting M unit basis of the image data record, and grouping the selected image data record into a first image paragraph;

determining content of the first image paragraph to generate an image content result, wherein the image content result comprises a dynamic content and a static content; and

detecting a change content for the image data record based on the image content result and generating the at least one image paragraph mark according to time position of the change content.

9. The segmentation system of claim 8, wherein wherein the image segmentation unit is further configured for executing following steps:

selecting T unit basis of the first image paragraph, and calculating similarity of the image data record in T unit basis to generate an image difference result;

if the image difference result is larger than a first image threshold, determining content the first image paragraph as the dynamic content; and

if the image difference result is less than the first image threshold, determining content the first image paragraph as the static content.

10. The segmentation system of claim 8, wherein the image segmentation unit is further configured for executing following steps:

if the first image paragraph is the dynamic content, calculating similarity between an Mth unit basis of the image data record and an (M+1)th unit basis of the image data record to generate an image difference value;

if the image difference value is larger than a second image threshold, merging the (M+1)th unit basis of the image data record into the first image paragraph; and

if the image difference value is less than the second image threshold, generating the at least one image paragraph mark in time position of the (M+1)th unit basis of the image data record, and selecting M unit basis of the image data record, and grouping the selected image data record into a second image paragraph.

11. The segmentation system of claim 8, wherein the image segmentation unit is further configured for executing following steps:

if the first image paragraph is the static content, calculating similarity between an Mth unit basis of the image data record and an (M+1)th unit basis of the image data record to generate an image difference value;

if the image difference value is less than a second image threshold, merging the (M+1)th unit basis of the image data record into the first image paragraph; and

if the image difference value is larger than the second image threshold, generating the at least one image paragraph mark in time position of the (M+1)th unit basis of the image data record, and selecting M unit basis of the image data record, and grouping the selected image data record into a second image paragraph.

12. The segmentation system of claim 7, wherein the audio segmentation unit is further configured for executing following steps:

transforming the audio data record into an audio time domain signal and an audio frequency domain signal, respectively;

selecting a time domain section from the audio time domain signal, and determining whether amplitude of the time domain section is less than a first threshold, if amplitude of the time domain section is less than the first threshold, generating the at least one audio paragraph mark; and

selecting a first frequency domain section and a second frequency domain section from the audio frequency domain signal, and determining whether a difference value of spectral intensity between the first frequency domain section and the second frequency domain section is greater than a second threshold, if the difference value of spectral intensity between the first frequency domain section and the second frequency domain section is larger than the second threshold, generating the at least one audio paragraph mark.

13. A non-transitory computer-readable medium including one or more sequences of instructions to be executed by a processor for performing a segmentation method, wherein the segmentation method comprises:

receiving a video data record, wherein the video data record comprises an image data record and an audio data record;

performing segmentation processing for the image data record to generate at least one image paragraph mark;

performing segmentation processing for the audio data record to generate at least one audio paragraph mark; and

calculating a time difference value between an image marking time of the at least one image paragraph mark and an audio marking time of the at least one audio paragraph mark and generating at least one video content mark,

wherein the time difference value is compared with a threshold to determine whether or not to use the image marking time and the audio marking time for generating the at least one video content mark, and

wherein when the time difference value is less than the threshold, the at least one video content mark is generated by the time average of the image marking time and the audio marking time.