

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
26 January 2012 (26.01.2012)

PCT

(10) International Publication Number  
**WO 2012/011735 A2**

(51) International Patent Classification:  
*H04N 7/173* (2011.01) *H04N 7/24* (2011.01)

(74) Agent: **LEE, Keon-Joo**; Mihwa Bldg. 110-2, Myongryun-dong 4-ga, Chongro-gu, Seoul 110-524 (KR).

(21) International Application Number:  
PCT/KR2011/005337

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

(22) International Filing Date:  
20 July 2011 (20.07.2011)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
10-2010-0070225 20 July 2010 (20.07.2010) KR  
10-2010-0073096 28 July 2010 (28.07.2010) KR  
10-2010-0097815 7 October 2010 (07.10.2010) KR

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(71) Applicant (for all designated States except US): **SAM-SUNG ELECTRONICS CO., LTD.** [KR/KR]; 416, Maetan-dong, Yeongtong-gu, Suwon-si, Gyeonggi-do 442-742 (KR).

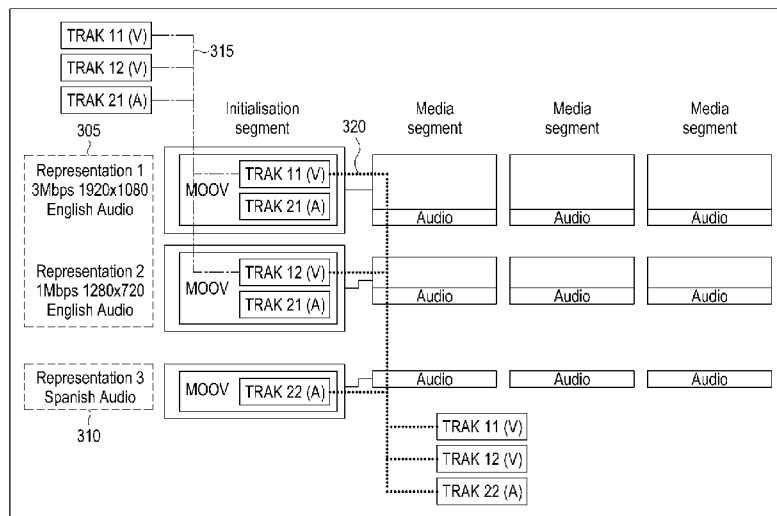
(72) Inventors: **RHYU, Sung-Ryeul**; #514-1204, Dongmun Goodmorning Hill, 5-cha APT., Dongcheon-dong, Suji-gu, Yongin-si, Gyeonggi-do 448-710 (KR). **HWANG, Seo-Young**; #101-907, Castle Tower, Ingye-dong, Paldal-gu, Suwon-si, Gyeonggi-do 442-070 (KR). **SONG, Jae-Yeon**; #B-805, Sungbo APT., Yeoksam-dong, Gangnam-gu, Seoul 135-080 (KR).

Published:

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

(54) Title: METHOD AND APPARATUS FOR TRANSMITTING AND RECEIVING ADAPTIVE STREAMING MECHANISM-BASED CONTENT

[Fig. 3]



(57) Abstract: Methods and apparatus are provided for transmitting adaptive streaming mechanism-based content. Representations, which include content components, are combined into a group. At least two of the group, a representation, and a content component type are combined into a subset. The subset is transmitted.

WO 2012/011735 A2

## Description

### **Title of Invention: METHOD AND APPARATUS FOR TRANSMITTING AND RECEIVING ADAPTIVE STREAMING MECHANISM-BASED CONTENT**

#### **Technical Field**

- [1] The present invention relates generally to an adaptive streaming service, and more particularly, to a method and apparatus for transmitting and receiving broadcast content using an adaptive streaming service.

#### **Background Art**

- [2] Recently, due to an increase in the use of on-line content, the importance of seamless content delivery has increased significantly. More attention is now concentrated on adaptive streaming technology in which a bit rate of content is smoothly or adaptively changed during its delivery depending on the change in available transmission bandwidth and/or performance of a user device. This adaptive streaming technology optimizes the bit rate by analyzing the network environment of the user. Multiple standards organizations have discussed new transmission technology like the adaptive streaming, in which a user terminal selects unit content, segmented into small pieces based on a Hyper Text Transfer Protocol (HTTP), and plays it depending on the situation, allowing its user to enjoy the content seamlessly and without delay.
- [3] Internet Protocol (IP)-based convergence of telecommunications and broadcasting, which is a marketable new next-generation business model, has attracted attention as one of the next-generation market-leading technologies. Particularly, Internet Protocol Television (IPTV), the largest pillar of Convergence of Broadcasting and Mobile Service (CBMS), is technology that can provide new services by recombining or combining the existing TV, voice, and data technologies. IPTV can also support interactive features, making it possible to create a variety of business models and services. Compared with the existing service provision technology specialized for limited terminal environments, this IPTV technology may provide its services to fixed-type terminals, such as set-top boxes, PCs, and TVs, and mobile-type terminals, such as cell phones and Personal Digital Assistants (PDAs) in both a wired network and a wireless network (or a broadcast network) regardless of the network type. Thus, it may be possible to provide and apply the same services in various different environments and enabling the service providers to develop a variety of new services taking various environments into consideration.

#### **Disclosure of Invention**

### **Technical Problem**

- [4] The present invention has been made to address at least the above problems and/or disadvantages and to provide at least the advantages described below. Accordingly, an aspect of the present invention provides a method and apparatus for transmitting and receiving adaptive streaming mechanism-based broadcast content in a digital broadcasting system.

### **Solution to Problem**

- [5] According to one aspect of the present invention, a method is provided for transmitting adaptive streaming mechanism-based content. Representations that include content components are combined into a group. At least two of the group, a representation, and a content component type are combined into a subset. The subset is transmitted.
- [6] According to another aspect of the present invention, an apparatus is provided for transmitting adaptive streaming mechanism-based content. The apparatus includes a controller for combining representations, which include content components, into a group, and combining at least two of the group, a representation, and a content component type into a subset. The apparatus also includes a transmitter for transmitting the subset.
- [7] According to an additional aspect of the present invention, a method is provided for receiving adaptive streaming mechanism-based content. A subset is received. The subset is configured by combining representations, which include content components, into a group and combining at least two of the group, a representation, and a content component type into the subset. At least one of a desired representation, a desired group, and a desired content component type is selected from the received subset, and the at least one selection is played.
- [8] According to a further aspect of the present invention, an apparatus is provided for receiving adaptive streaming mechanism-based content. The apparatus includes a receiver for receiving a subset. The subset is configured by combining representations, which include content components into a group and combining at least two of the group, a representation, and a content component type into the subset. The apparatus also includes a controller for selecting at least one of a desired representation, a desired group, and a desired content component type from the received subset, and playing the at least one selection.

### **Brief Description of Drawings**

- [9] The above and other aspects, features and advantages of the present invention will be more apparent from the following detailed description when taken in conjunction with the accompanying drawings, in which:

- [10] FIG. 1 is a diagram illustrating possible problems of initialization segments during storage of content in a 3<sup>rd</sup> Generation Partnership Project (3GPP) HyperText Transfer Protocol (HTTP) Adaptive Streaming (HAS) system to which an embodiment of the present invention is applicable;
- [11] FIG. 2 is a diagram illustrating initialization segments reconfigured taking their storage into account, according to an embodiment of the present invention;
- [12] FIG. 3 is a diagram illustrating an example of service combinations available in multiple representations, according to an embodiment of the present invention; and
- [13] FIG. 4 is a diagram illustrating an example of configuring a segment to be available in a Content on Demand (CoD) server after a time timeShiftBufferDepth a Multimedia Presentation Description (MPD) uses to indicate a valid period of the segment of Live content, according to an embodiment of the present invention.

### **Mode for the Invention**

- [14] Embodiments of the present invention are described in detail with reference to the accompanying drawings. The same or similar components may be designated by the same or similar reference numerals although they are illustrated in different drawings. Detailed descriptions of constructions or processes known in the art may be omitted to avoid obscuring the subject matter of the present invention.
- [15] Although the names of entities defined in Open IPTV Forum (OIPF) will be used intact in the following description for convenience purpose only, they are not used to limit the scope of the present invention, and it will be understood by those of ordinary skill in the art that the present invention may be applied to any other systems having similar technical backgrounds.
- [16] Prior to a description of the present invention, the terms of 3GPP HAS, used in this specification, are defined below.
- [17] HAS: HTTP Adaptive Streaming
- [18] Content: creations such as movies, music, and news, and a storage form thereof
- [19] Content component: components (e.g., video, audio, caption, etc.) constituting content
- [20] Representation: one representation form. Low-quality and high-quality representations may exist, but their contexts are the same.
- [21] 3GPP HAS includes representations created by encoding the same content and content components with different conditions (i.e., different network bandwidths, resolutions, codecs, etc.), and an MPD, serving as a manifest file, configured by dividing the full length into several equivalent periods, re-dividing the periods into segments of a sufficiently short time of about 1 to 10 seconds (segmentation), and specifying file names, sequence numbers, start times, lengths, etc. of the segments.

- [22] The segment includes an initialization segment having information for decoder configuration provided to properly decode its associated representation, and media segments including segmented content components. A client first receives an MPD and checks which representation is included in it. Thereafter, the client determines performance and capability of the client device, a network speed, etc., selects a proper representation, downloads fragments thereof in sequence, and starts playing them.
- [23] If there is a change in mutable variables such as a network bandwidth, i.e., if the network speed increases, the client selects fragments of the representation created with a higher bandwidth, and plays the selected fragments. As a result, while the image quality may vary depending on the network speed, the playing may continue without being disconnected.
- [24] The requirements for the 3GPP HAS technology having the above characteristics, the problems thereof, and solutions thereto (i.e., embodiments of the present invention) are described in detail below.
- [25] With respect to 3GPP HAS content, proper segments may be selected from among the segments in multiple representations depending on conditions such as the network speed, and then played. The content can be stored by concatenating the selected segments in sequence. However, the representations may have different initialization segments, because they are created with different encoding options.
- [26] FIG. 1 is a diagram that illustrates possible problems of initialization segments during storage of content in a 3GPP HAS system to which an embodiment of the present invention is applicable.
- [27] Referring to FIG. 1, representations have different initialization segments as shown by reference numeral 100. In attaching received media segments 110 thereto for their storage, storing only one of the existing initialization segments is equivalent to losing the information used to play media segments of other presentations.
- [28] Accordingly, a first embodiment of the present invention provides methods for solving the possible problems which may occur when storing the 3GPP HAS content.
- [29] The first embodiment of the present invention may store media segments of different representations so as to be later replayed, in the following two ways.
- [30] In a first case, a client device may have an ISOFF reconfiguration feature.
- [31] In this case, the client device may reconfigure initialization segments of a new version through the following steps.
- [32] - All initialization segments, or an initialization segment corresponding only to the representation, which is downloaded depending on conditions such as the network speed, are concatenated or merged in the following order;
- [33] - Each initialization segment includes boxes such as 'styp' indicating that its type is a segment, and 'moov';

- [34] - 'trak' boxes are gathered in their 'moov' boxes to be merged;
- [35] - track IDs of 'trak' boxes are substituted with new IDs if they overlap; and
- [36] - Initialization segments are reconfigured and stored using 'moov' boxes, which are newly made by putting all 'trak' boxes in the 'moov' boxes.
- [37] If media segments are stored intact together with the reconfigured initialization segments, the following problems may occur. Specifically, 1) if track IDs overlap, track IDs may not be present in associated media segments because track IDs of 'trak' boxes in 'moov' boxes of the initialization segments were substituted at random, or 2) if track IDs of media segments are the same, the client device may not normally replay them since proper decoding configuration may not be applied thereto.
- [38] Therefore, the client device should reconfigure even media segments of a new version in accordance with the following steps.
- [39] - After downloading media segments, the client device should change their track IDs depending on the substitution rule;
- [40] - Each media segment includes boxes such as 'styp' indicating that its type is a segment, 'sidx', 'tfad', 'moof', 'mdat';
- [41] - If there is a 'sidx' box for substitution, the client device opens it and substitutes its reference track ID;
- [42] - For substitution, the client device opens a 'traf' box in a 'moof' box and substitutes its track ID; and
- [43] - The client device stores the media segments using the modified 'sidx' and 'moof' boxes.
- [44] If continuously stored or concatenated, the reconfigured initialization segments and media segments may be replayed later.
- [45] In a second case of the first embodiment, a client provider may consider storing content.
- [46] If the content provider provides initialization segments considering the possible situations where content may be stored in a client device, the client device may store segments intact without modification and replay them later.
- [47] The content provider configures track IDs for all representations so as not to overlap.
- [48] The content provider configures an initialization segment of each representation to have decoding configurations for content components in all other representations.
- [49] In addition, the content provider may configure an MPD so that all representations may indicate the same initialization segment.
- [50] FIG. 2 illustrates initialization segments reconfigured according to the first embodiment of the present invention.
- [51] Accordingly, in 3GPP HAS, the following changes are required for an MPD and a file format.

- [52] - MPD:
- [53] - All representations indicate the same initialization segment file;
- [54] - oipf:component is used to describe which content component is played when a representation is selected using the common initialization segment of the representations. It is used to further describe IDs and types of content components of each representation; and
- [55] - All oipf:components uses an existing optional ID attribute.
- [56] - File Format:
- [57] - Initialization segments include information for decoding tracks of all representations. Therefore, all 'trak' boxes may be included in one 'moov' box; and
- [58] - track IDs are be unique; track IDs of different content components do not overlap.
- [59] 3GPP HAS content may have configurations of combinations such as 'video+audio', 'video without audio', and 'audio without video', as representations are combined.
- [60] However, in the case where a client device makes its best effort depending on conditions such as the network speed, a content creator has no way to explicitly describe the combinations.
- [61] FIG. 3 illustrates an example of service combinations available in multiple representations, according to a second embodiment of the present invention.
- [62] Referring to FIG. 3, when representations 305 exist, which are created in English and have different bit rates, and a Spanish audio track 310 exists separately, a content creator may enable several services such as 'video + English audio'315 and 'video + Spanish audio'320, by combining content components.
- [63] Accordingly, the second embodiment of the present invention provides a method in which a content creator may combine content components to enable several services.
- [64] In order to represent that a content creator may enable several services by combining content components, the second embodiment of the present invention defining services indicating combinations of a representation, content components in the representation, and groups (collections of representations), as shown in Table 1 below.
- [65] Table 1

[Table 1]

```

<Services>
  <service ID="1" description="Movie (non-adaptive English version)">
    <representationInfo refID="1"/> // high quality representation
  </service>
  <service ID="2" description="Movie (adaptive English version)">
    <groupInfo refID="1"/> //alternateGroup of English Movie
  </service>
  <service ID="3" description="Movie (adaptive Spanish version)">
    <groupInfo refID="1" type="video"> //alternateGroup of video
    <representationInfo refID="3"/> // Spanish audio
  </service>
  <service ID="4" description="Movie (non-adaptive Spanish version)">
    <componentInfo refID="11"/> // high quality video
    <representationInfo refID="3"/> // Spanish audio
  </service>
</Services>
<Representation ID="1" group="1" bandwidth="6Mbps" ...>
  <SegmentInfo ...>
    ...
  <SegmentInfo>
    <oipf:Components>
      <oipf:component type="video" id="11">
      <oipf:component type="audio" id="21" >
    </oipf:Components>
  </Representation>
<Representation ID="2" group="1" bandwidth="3Mbps" ...>
  <SegmentInfo ...>
    ...
  <SegmentInfo>
    <oipf:Components>
      <oipf:component type="video" id="12">
      <oipf:component type="audio" id="21" >
    </oipf:Components>
  </Representation>
<Representation ID="3" group="2" bandwidth="1Mbps" ...>
  <SegmentInfo ...>
    ...
  <SegmentInfo>
    <oipf:Components>
      <oipf:component type="audio" id="22" >
    </oipf:Components>
  </Representation>

```

[66] In Table 1, 'Services' represents the presence of combinable services, and 'Services' represents each of the services. In Service, representation, group, content component, etc. are referenced, and if a type is specified for a group, it represents that only one

content component type such as video and audio in the group is referenced.

[67] Accordingly, the following changes are required in 3GPP HAS.

[68] - MPD:

[69] - 'Period' has services as its elements;

[70] - 'Services' has a service(s) as its element;

[71] - 'Service' has an ID and a description as its attributes;

[72] - 'Service' has representationInfo, groupInfo, componentInfo as its elements;

[73] - 'RepresentationInfo' indicates a representation to be referenced by using refID as an attribute;

[74] - 'GroupInfo' indicates a group to be referenced by using refID as an attribute;

[75] - 'GroupInfo' indicates a content component type to be referenced in a group by using 'type' as an attribute;

[76] - 'Representation' has ID added as an attribute; and

[77] - If service elements are present, all representations should have their IDs.

[78] This service information may exist in multiple locations such as an MPD or another MPD, and another segment file.

[79] 3GPP HAS uses an MPD for content delivery, but does not describe conversion from the stored format to the MPD. Therefore, all delivery-related information is in the MPD. There is no information for conversion to the delivery format if the MPD is not stored during storage. In addition, 3GPP HAS embeds an MPD or a link indicating the MPD in initialization segments during storage, but this is not mandatory.

[80] Accordingly, a third embodiment of the present invention provides the following method of forcing an MPD to be embedded in initialization segments and performing conversion to the delivery format if the MPD is present.

[81] - For conversion to the delivery format, information about representations is read from the MPD, and track information is read from a stored file;

[82] - The stored file is divided on a segment basis depending on the track information, and stored in a file; and

[83] - If the delivery format corresponds to a self-initialization segment, the MPD is updated by calculating an offset from its start point.

[84] In the third embodiment of the present invention, the following change is required in 3GPP HAS.

[85] - MPD:

[86] - Initialization segments should have an MPD and its link embedded therein.

[87] 3GPP HAS does not describe whether to store previously transmitted content components, a storage period thereof, and/or a method of indicating stored content components. In 3GPP HAS, a start attribute value specified in Period is applied to the entire representation (or a set of segments) in the Period. Representations have no

separate time information, and segments have only duration values. When content is cached, it may be cached over the full Period, meaning that content components may not be used independently, or may not be reused.

- [88] Accordingly, in a fourth embodiment of the present invention, in order to play the previously transmitted content and the current content together, content components should be stored such that they may be reused later after a previous transmission. In addition, a duration or period should be specified, which indicates when the content components will be reused, and indicates whether they will be reused since then. In addition, components previously transmitted and stored in the local are instructed to be played together with the current components, and if no components are stored, an available link should be provided to a server as shown in Table 2 below.

- [89] Table 2

[Table 2]

```
<Representation expirationUTC="2010-08-31" identifier="ABC123">
  <segmentInfo baseUrl="rep2/">
    <Url sourceUri="seg1.3gp">
  </segmentInfo>
</Representation>
```

- [90] In Table 2, 'expirationUTC' represents an absolute time and indicates when this representation will be reused, and 'Identifier' is used as an identifier for distinguishing a stored representation.
- [91] Based on the above attributes, the client device may manage reusing the previously transmitted content components in the following order, storing the content components to reuse them later, and/or deleting expired content components.
- [92] - In an MPD, expirationUTC is presented in Representation, Content component, Group, etc. as an attribute;
- [93] - Using Identifier or URL of a segment as an identifier, the client device determines if the associated content components or individual files are stored in a storage unit connected to the client device;
- [94] - The client device includes a step of determining whether an identifier is the same and a segment file's name is the same, or whether a part of a URL of a segment is equal to a part of a path of a stored file, or whether all or part of a file managed by a separate database is the same;
- [95] - If all or part of a stored file is discovered, the client device plays the file;
- [96] - If there is no stored file, a relevant file or a part of the file is received from a content server and then stored in a storage unit connected to the client device;
- [97] - For storage, Identifier and/or base URL of a segment may be used as a part of a

folder name or a file name;

- [98] - To describe a storage period, a separate database may be used, or a storage period may be used as a folder name or a part of the folder name; and
- [99] - If the storage unit connected to the client device lacks empty space at the moment a new content component is to be stored, the client device deletes files whose valid period has expired, deletes files whose valid period has expired by searching the storage unit upon start up the client device, or deletes files corresponding to the time the valid period expires, by checking the database.
- [100] Accordingly, in the fourth embodiment of the present invention, the following changes are required in 3GPP HAS.
- [101] - MPD:
- [102] - When later replayed, representation should have expirationUTC and Identifier;
- [103] - ExpirationUTC indicates the time the file stored to be later replayed may be deleted as it is no longer used; and
- [104] - Identifier represents an identifier used to reuse the same content component in different MPDs.
- [105] 'Timed metadata' is meaningful at a specific time or for a specific period from the time. 3GPP HAS content is set to display its information such as a title and a copyright, in ProgramInformation.
- [106] However, ProgramInformation, or metadata corresponding to the entire MPD, is not timed metadata. In addition, it is not possible to specify the time the metadata is applied, or to specify a time period for which the metadata is applied.
- [107] Accordingly, a fifth embodiment of the present invention provides a method for applying timed metadata to both the MPD and the File Format. Since the MPD describes one or more Periods, in order to propose a time of timed metadata, Timed Metadata is newly established, and its start time, duration and metadata are described in Table 3 below. The Timed Metadata, an element capable of expanding content features, can realize a variety of effects by being inserted in various locations depending on the scene and the content features. For example, the Timed Metadata may describe media by being applied to a specific time in the time domain after being applied to a Period level. In addition, the Timed Metadata may realize other features depending on media attributes such as audio and video, by being included in a Representation level. The Timed Metadata may be included in a container containing actual content data except for the MPD. The features the Timed Metadata desires to realize may be described in the text, and for the features, a URL or a location in a media representation file the MPD desires to realize, may be referred to.
- [108] Table 3

[Table 3]

<pre> &lt;Period&gt;   &lt;TimedMetadata startTime="00:00:10" duration="00:02:00"&gt;     Car chase scene   &lt;/TimedMetadata&gt; &lt;/Representation&gt; ... </pre>
---

[109] Table 4 is an example in which timedMetadata is included in an MPD on a representation level, according to an embodiment of the present invention.

[110] Table 4

[Table 4]

<pre> ... &lt;Period start="PT0s"&gt;   &lt;Representation mimeType="video/iso6092"&gt;     &lt;timedMetadata startTime="PT10s" duration="PT60s"&gt;       Car chase scene     &lt;/timedMetadata&gt;&lt;SegmentInfo duration="PT120s" baseUrl="rep2"/&gt;     &lt;InitialisationSegmentURL sourceURL="seg-init.3gp"/&gt;     &lt;Url sourceURL="seg-1.3gp"/&gt;     &lt;Url sourceURL="seg-2.3gp"/&gt;     &lt;Url sourceURL="seg-3.3gp"/&gt;   &lt;/SegmentInfo&gt; &lt;/Representation&gt; &lt;Representation mimeType="audio/aac"&gt;   &lt;timedMetadata startTime="PT10s" duration="PT60s" sourceURL="meta.xml"/&gt;   &lt;SegmentInfo duration="PT120s" baseUrl="rep2"/&gt;   &lt;InitialisationSegmentURL sourceURL="seg-init.3gp"/&gt;   &lt;Url sourceURL="seg-1.3gp"/&gt;   &lt;Url sourceURL="seg-2.3gp"/&gt;   &lt;Url sourceURL="seg-3.3gp"/&gt; &lt;/SegmentInfo&gt; &lt;/Representation&gt; ... &lt;/Period&gt; ... </pre>
--

[111] File Format proposes a new box as shown in Table 5 below, capable of inserting metadata in units of movie fragment in a segment. Table 6 is a description for Table 5.

[112] Table 5

[Table 5]

<i>The Movie Fragment Meta box</i>	
<b>Definition</b>	
Box Type:	'mfmt'
Container:	File, Movie Fragment Box ('moof'), Track Fragment Box ('traf')
Mandatory:	No
Quantity:	Zero or one
<b>Syntax</b>	
<pre>aligned(8) class MovieFragmentMetaBox (handler_type)   extends FullBox('mfmt', version = 0, 0) {     HandlerBox(handler_type)  theHandler;   } aligned(8) class TimedMetaInfo extends FullBox('tmif', version, 0) {   unsigned int(32)  entry_count;   for (i=1; i&lt;=entry_count; i++) {     if (version == 0) {       unsigned int(32) item_ID;       unsigned int(32) reference_track_ID;       int(32) mediaTime;       unsigned int(32) duration;       string          item_name;       string          content_type;       string          content_encoding;  //optional     }     else { // version == 1       unsigned int(64) item_ID;       unsigned int(64) reference_track_ID;       int(64) mediaTime;       unsigned int(64) duration;       string          item_name;       string          content_type;       string          content_encoding;  //optional     }   } }</pre>	

[113] Table 6

[Table 6]

<b>Semantics</b>
The structure or format of the metadata is declared by the handler
entry_count counts the number of resources of timed metadata in the following array.
item_ID is an arbitrary integer 'name' for this resource which can be used to refer to it (e.g. in a URL).
reference_track_ID provides the track_ID for the reference track
media_time is an integer containing the starting time of this timed metadata is present (composition time)
duration is an integer that declares length of the presentation (in the indicated timescale).
item_name is a null-terminated string in UTF-8 characters containing a symbolic name of the item.
content_type is a null-terminated string in UTF-8 characters with the MIME type of the item. If the item is content encoded (see below), then the content type refers to the item after content decoding.
content_encoding is an optional null-terminated string in UTF-8 characters used to indicate that the binary file is encoded and needs to be decoded before interpreted. The values are as defined for Content-Encoding for HTTP/1.1. Some possible values are "gzip", "compress" and "deflate". An empty string indicates no content encoding. Note that the item is stored after the content encoding has been applied.

[114] 3GPP HAS realizes Live streaming by splitting content on a period basis and a segment basis, providing an MPD describing the segments, and allowing the client device to constantly update the MPD at regular intervals and transmit segment files of the next play period. However, when the client device pauses play, since a valid period of the segments the MPD currently describes is bound with the valid period of the MPD, if the MPD is updated with new information, the previously described segments are no longer available in the server. In addition, when the client device attempts time-shifted play back to the previous time, if segment information for the time is not present in the MPD, the client device may not know a name and availability of the segment file associated with the time. In addition, a server for delivering Live content does not guarantee storage of past segments except for the current information in order to ensure real-time performance. However, a CoD server is mainly used to provide more content than real-time content.

[115] FIG. 4 illustrates an example of configuring a segment to be available in a CoD server after a time `timeShiftBufferDepth` an MPD uses to indicate a valid period of the segment of Live content, according to a sixth embodiment of the present invention.

[116] Referring to FIG. 4, if periods are specified separately and segments associated with the periods are transmitted by Live and CoD servers, a period 405 may be received from the Live server and a period 410 may be received from the CoD server.

[117] The period 405 of Live content shifts over time.

[118] If a play time shifts to the period 410 by Pause or time-shift, the client device receives segment files using a URL for the CoD server.

[119] An MPD is configured as shown in Table 7 below.

[120] Table 7

[Table 7]

```
<MPD timeshiftbufferdepth="PT5S"
  availableStartTime="2010-08-31"
  timeshiftUTC="2010-08-31"
  baseUrl="HTTP://www.foo.com/live"
  basePvrUrl="HTTP://www.foo.com/cod">
```

[121] Accordingly, in the sixth embodiment of the present invention, the following changes are required in 3GPP HAS.

[122] - MPD:

[123] - `timeshiftUTC` and `basePvrUrl` are added in the MPD as attributes;

[124] - `timeshiftUTC` indicates content from which time in a program is available on the CoD server; and

[125] - `basePvrURL` indicates a base URL used to receive segments from the CoD server.

- [126] 3GPP HAS may make random access to CoD content using MFRA of ISOFF, but Live content has no information for random access. In addition, 3GPP HAS has proposed a 'sidx' box for segment-continued play during switching of play representations. However, the 'sidx' box has insufficient information for random access as it has information about only one random access point per one movie fragment (normally, having 2 ~ 3 seconds).
- [127] Accordingly, a seventh embodiment of the present invention provides a method with the following steps.
- [128] - The 'sidx' box is expanded to indicate all Random Access Points (RAPs) in movie fragments;
- [129] - A start point and an end point of a sample corresponding to a RAP are specified to make it possible to receive the RAP by request without parsing a 'moof' box; and
- [130] - Due to the change, the client device may start play beginning from a more accurate random access point after parsing the 'sidx' box.
- [131] The client device may use the 'sidx' box for trick play. To this end, the following method steps are provided.
- [132] - The client device receives the 'sidx' box from a segment of a representation it will play;
- [133] - The client device receives a random access point list for the 'sidx' box, and determines which random access point it will play, depending on the speed of a trick mode; and
- [134] - The client device receives and plays only the necessary samples using the start point and end point of samples of the determined random access points.
- [135] Accordingly, in the seventh embodiment of the present invention, the following changes of Table 8 below are required in 3GPP HAS.
- [136] Table 8

[Table 8]

```

Syntax of 'sidx':
    aligned(8) class SegmentIndexBox extends FullBox('sidx', version, 0) {
        unsigned int(32) reference_track_ID;
        unsigned int(16) track_count;
        unsigned int(16) reference_count;
        for (i=1; i <= track_count; i++)
        {
            unsigned int(32) track_ID;
            if (version==0)
            {
                unsigned int(32) decoding_time;
            } else
            {
                unsigned int(64) decoding_time;
            }
        }
        for(i=1; i <= reference_count; i++)
        {
            bit (1)          reference_type;
            unsigned int(31) reference_offset;
            unsigned int(32) subsegment_duration;
            bit(1)  contains_RAP;
            unsigned int(31) rap_count; //added
            for(j=1;j<=rap_count;j++) //added
            {
                unsigned(32)  RAP_delta_time; //moved
                unsigned(32)  Sample_start_offset; //added
                unsigned(32)  Sample_end_offset; //added
            }
        }
    }
}

```

- [137] In 3GPP HAS, Live describes only a part of the entire content, as a part of an MPD is updated at regular intervals. However, in case of CoD, as it contains information about all parts of content, the client device receives unnecessary information to determine one segment required to start play.
- [138] Accordingly, an eighth embodiment of the present invention provides the following methods.
- [139] A method is provided for distributing information in an MPD and indicating them in accordance with the steps below.
- [140] - The MPD includes an MPD describing a Period including one or small segment containing information for initial execution, and one or more MPDs obtained by dividing the following time into periods;

- [141] - The MPD is embedded in a separate file or segment (MPD segment);
- [142] - The MPD segment is a media segment having an MPD in an 'mfmt' box; and
- [143] - Each MPD indicates a link of the next MPD, or a link of an embedded segment.
- [144] Accordingly, in the eighth embodiment of the present invention, the following changes are required in 3GPP HAS.
- [145] - MPD:
- [146] - MPDSegmentURL is added in SegmentInfo as an element; and
- [147] - MPDSegmentURL has sourceURL and range as its attributes.
- [148] - File Format:
- [149] - The MPD segment is a media segment having an MPD in an 'mtmt' box.
- [150] Although bandwidth, width, height, lang, mimetype, qualityranking, etc. have been considered for interactive selection, views have not been considered. Views of multi-views may be different only in location and number on the condition that different attributes such as bandwidth have the same value.
- [151] As for multi-view content, view selection is impossible because the client device may not determine which view it receives, based on only the information provided from the current representation.
- [152] There are many different types of relationships among views.
- [153] For multi-views, view corresponds to a relationship that at least one of them may be selected at random and displayed.
- [154] In 3D stereoscopic, Left and Right Views correspond to a relationship that they should be played together.
- [155] In case of SVC (Scalable Video Coding), if there is a layer relationship such as a base layer and an enhanced layer, it corresponds to a relationship that the base layer is selected, and then, the other enhanced layer may be selected in sequence.
- [156] Therefore, an attribute name such as relationship is proper, compared with view.
- [157] In a first method, an explicit description of relationships is provided below.
- [158] - Relationships:
- [159] - relationship="stereoscopic/L"
- [160] - relationship="stereoscopic/R" or
- [161] - relationship="multiview/1"
- [162] - relationship="multiview/2"
- [163] - relationship="multiview/3" or
- [164] - relationship="svc/base"
- [165] - relationship="svc/enhanced1"
- [166] - relationship="svc/enhanced2"
- [167] In a second method, a description of connections associated with relationships is provided below.

[168] - Relationships:

[169] - relationship="one or more" or

[170] - relationship="all" or

[171] - relationship="incremental/1"

[172] Accordingly, in a ninth embodiment of the present invention, the following changes are required in 3GPP HAS.

[173] - MPD:

[174] - Representation has relationship as an attribute; and

[175] - Relationship is configured in accordance with Method 1) and Method 2).

[176] In 3GPP HAS, since representation is "one of the alternative choices of the media content", information providing the selection criteria for choices includes the following attributes of representations:

[177] 1) Bandwidth: Bandwidth attribute

[178] 2) Display size: Width and height attribute

[179] 3) Language: lang attribute

[180] 4) Prioritization: QualityRanking attribute

[181] However, there is no proposed method capable of determining priorities of Representations with a QualityRanking attribute and Representations without a QualityRanking attribute.

[182] Accordingly, a tenth embodiment of the present invention provides a method having the following steps.

[183] - For alternate selection in a group, QualityRanking is inserted into all Representations, or not inserted into all Representations;

[184] - For priorities of groups, a low group id is indicated to have high priority; and

[185] - These priorities may be represented by separate attributes.

[186] 3GPP HAS provides attributes for describing differences between representations, but it does not indicate initial selection values and default values.

[187] Accordingly, an eleventh embodiment of the present invention provides a method having the following steps.

[188] - A value indicating initial selection is attached to one of Representations having a relationship of AlternateGroup;

[189] - Default is added as an attribute of Representation;

[190] - A value indicating initial selection is attached to one of Services; and

[191] - Default is added as an attribute of Service.

[192] As is apparent from the foregoing description, according to embodiments of the present invention, streaming broadcast content may be adaptively delivered in a digital broadcasting system.

[193] While the invention has been shown and described with reference to certain em-

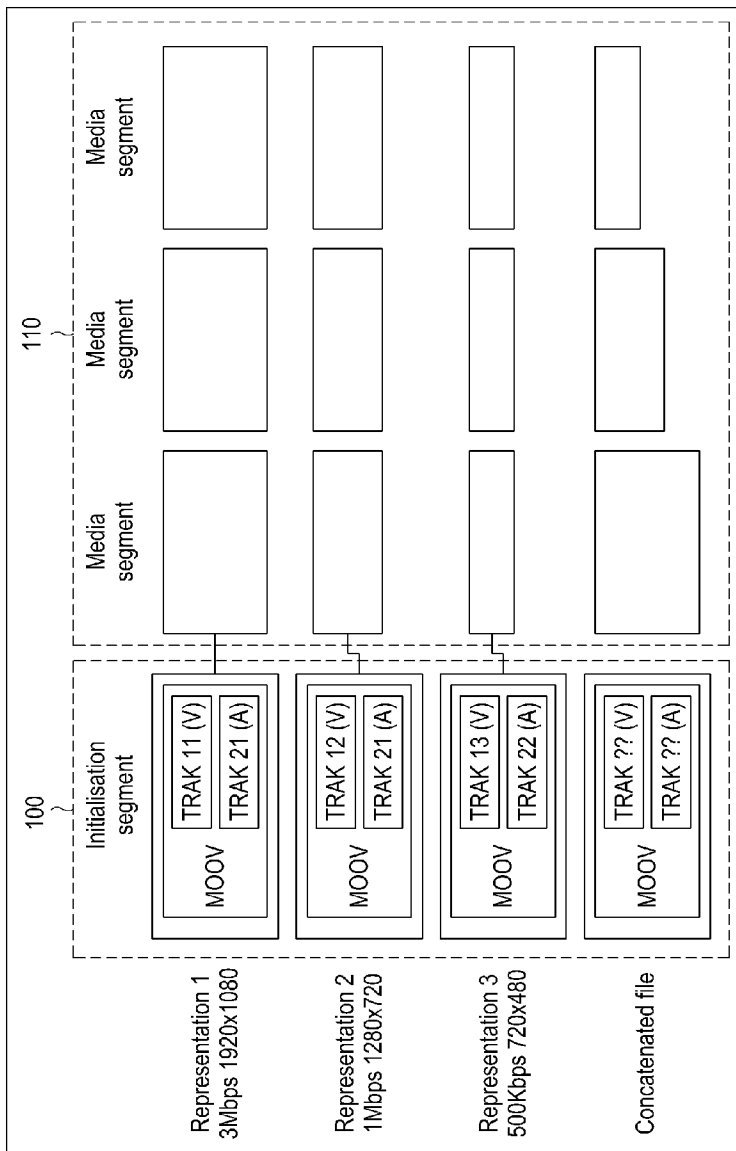
bodiments thereof, it will be understood by those skilled in the art that various changes in form and detail may be made therein without departing from the spirit and scope of the invention as defined by the appended claims and their equivalents.

## Claims

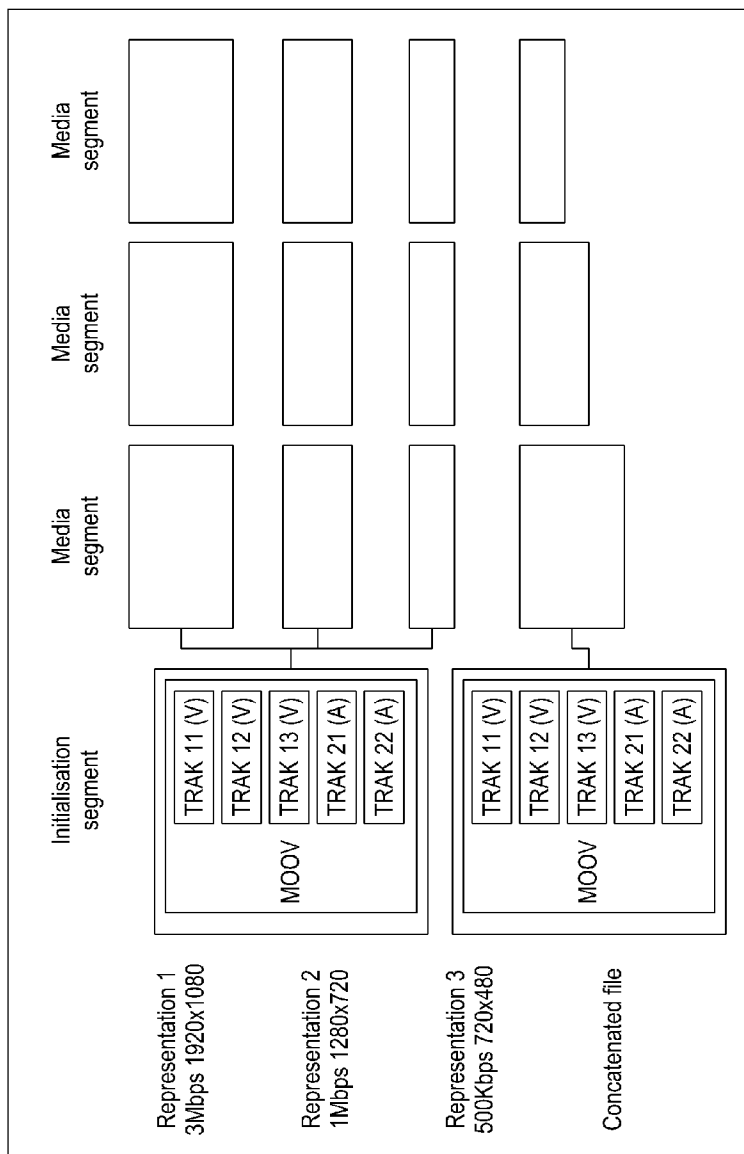
- [Claim 1] A method for transmitting adaptive streaming mechanism-based content, comprising the steps of:  
combining representations, which include content components, into a group;  
combining at least two of the group, a representation, and a content component type into a subset; and  
transmitting the subset.
- [Claim 2] An apparatus for transmitting adaptive streaming mechanism-based content, comprising:  
a controller for combining representations, which include content components, into a group, and combining at least two of the group, a representation, and a content component type into a subset; and  
a transmitter for transmitting the subset.
- [Claim 3] A method for receiving adaptive streaming mechanism-based content, comprising the steps of:  
receiving a subset, wherein the subset is configured by combining representations, which include content components, into a group, and combining at least two of the group, a representation, and a content component type into the subset; and  
selecting at least one of a desired representation, a desired group, and a desired content component type from the subset, and playing the at least one selection.
- [Claim 4] An apparatus for receiving adaptive streaming mechanism-based content, comprising:  
a receiver for receiving a subset, wherein the subset is configured by combining representations, which include content components, into a group, and combining at least two of the group, a representation, and a content component type into the subset; and  
a controller for selecting at least one of a desired representation, a desired group, and a desired content component type from the subset, and playing the at least one selection.
- [Claim 5] The method of claim 1 or claim 3 or the apparatus of claim 2 or claim 4, wherein the subset is a service.
- [Claim 6] The method of claim 1 or claim 3 or the apparatus of claim 2 or claim 4, wherein the subset is disposed in a Multimedia Presentation Description (MPD).

- [Claim 7] The method of claim 1 or claim 3 or the apparatus of claim 2 or claim 4, wherein the subset is an element of a period.
- [Claim 8] The method or apparatus of claim 5, wherein the service indicates at least one of representation information, group information and content component type information as a component element thereof.
- [Claim 9] The method or apparatus of claim 8, wherein the representation information indicates a representation that is to be referenced when a reference ID (refID) is used as an attribute.
- [Claim 10] The method or apparatus of claim 8, wherein the group information indicates a group that is to be referenced when a reference ID (refID) is used as an attribute.
- [Claim 11] The method or apparatus of claim 8, wherein the content component type information indicates a content component type that is to be referenced when a type is used as an attribute.

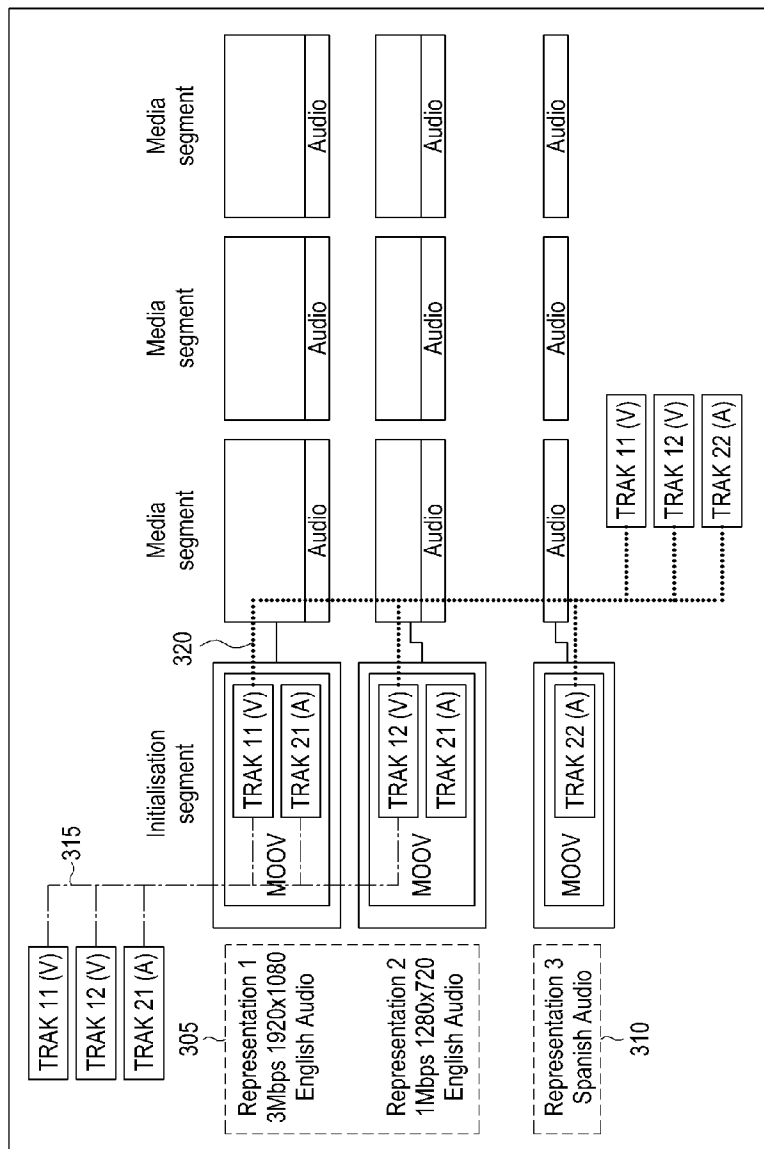
[Fig. 1]



[Fig. 2]



[Fig. 3]



[Fig. 4]

