

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
15 September 2011 (15.09.2011)

PCT

(10) International Publication Number  
**WO 2011/111987 A2**

(51) International Patent Classification:  
H04N 7/173 (2011.01) H04N 7/08 (2006.01)

(21) International Application Number:  
PCT/KR2011/001604

(22) International Filing Date:  
8 March 2011 (08.03.2011)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
10-2010-0020544 8 March 2010 (08.03.2010) KR

(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: **PARK, Kyung-Mo**; #23-1301, Eunma APT., Daechi 2-dong, Gangnam-gu, Seoul 135-969 (KR). **SONG, Jae-Yeon**; #B-805, Sungbo APT., Yeoksam-dong, Gangnam-gu, Seoul 135-080 (KR). **RHYU, Sung-Ryeul**; #514-1204, Dongmun Goodmorning Hill 5-cha APT., Dongcheon-dong, Suji-gu, Yongin-si, Gyeonggi-do 448-710 (KR).

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

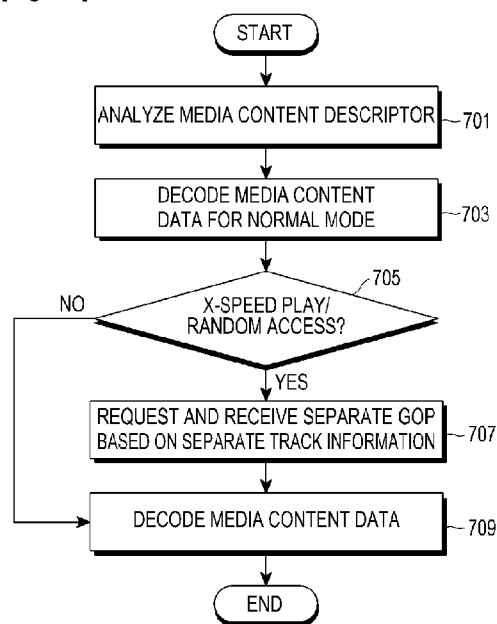
(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, 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.

(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).

[Continued on next page]

(54) Title: APPARATUS AND METHOD FOR PLAYING MEDIA CONTENT DATA

[Fig. 10]



(57) Abstract: A method is provided for playing media content by a client in a data communication network, in which a request for a media content descriptor is sent to a server, a media content descriptor including a Uniform Resource Locator (URL) for an access to specific media content data configured for execution of a specific function is received from the server, and the specific media content data is received using the URL included in the received media content descriptor.

WO 2011/111987 A2

**Published:**

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

## Description

### Title of Invention: APPARATUS AND METHOD FOR PLAYING MEDIA CONTENT DATA

#### Technical Field

- [1] The present invention relates generally to an apparatus and method for playing media content data, and more particularly, to a media content playing apparatus and method for minimizing delays in an X-speed play function (where X is a specific number) and a random access function regarding play of media content data received from a data communication network.

#### Background Art

- [2] In general, methods for accessing a variety of media content services over a data communication network may be roughly classified into a download method and a streaming method. The media content services refer to services provided by various types of digital media data, such as broadcasts, movies, and music. Herein, "digital media data" refers to digital information constituting associated media content, such as videos, sounds, scenes, etc.
- [3] The download method allows a user to download digital media data constituting desired media content service in non-real time, store the downloaded digital media data in a storage media, and play the stored digital media data at a desired time. The download method is mainly used for non-real time media content services, such as movies and dramas.
- [4] The streaming method allows a user to receive digital media data presently required, and play the received digital media data in real time. The streaming method is mainly used for real-time media content services corresponding to live broadcasts, such as news and sports broadcasts.
- [5] While receiving a media content service, a user may sometimes desire to play media content data at a faster playback speed than the normal playback speed (which is called 'X-speed play'), or to play a specific part of the media other than a part of the media currently being played (which is called 'random access').

#### Disclosure of Invention

##### Technical Problem

- [6] Conventionally, if a user requests X-speed play or random access based on a Moving Picture Experts Group (MPEG) file format, the X-speed play or random access requested by the user may not be precisely performed, because the existing MPEG file format is implemented based on a file format, in which broadcast by a server is considered, rather than a communication network between a client and a server is

considered.

- [7] Therefore, when the communication network between a client and a server is considered, it is not possible to precisely provide the X-speed play function and random access function required by the user based on the existing MPEG file format. Accordingly, there has been a long demand for a method for precisely providing the X-speed play and random access functions based on the existing MPEG file format.

### **Solution to Problem**

- [8] An aspect of the present invention is to address at least the above-mentioned problems and/or disadvantages and to provide at least the advantages described below. Accordingly, an aspect of the present invention is to provide a media content playing apparatus and method for precisely controlling streaming/download-based media content data.
- [9] Another aspect of the present invention provides a media content playing apparatus and method for controlling media content data consisting of stream groups for an X-speed play function and a random access function.
- [10] Another aspect of the present invention provides a media content playing apparatus and method for providing media content data capable of being adjusted and controlled according to various content types and user preferences by providing semantic configuration and role information of media content provided.
- [11] In accordance with one aspect of the present invention, there is provided a method for playing media content by a client in a data communication network. The method includes sending, to a server, a request for a media content descriptor; receiving, from the server, a media content descriptor including a Uniform Resource Locator (URL) for an access to specific media content data configured for execution of a specific function; and receiving the specific media content data using the URL included in the received media content descriptor.
- [12] In accordance with another aspect of the present invention, there is provided a method for supporting media content play of a client by a server in a data communication network. The method includes receiving, from the client, a request for a media content descriptor; and transmitting, to the client, a media content descriptor including a Uniform Resource Locator (URL) for an access to specific media content data configured for execution of a specific function.
- [13] In accordance with another aspect of the present invention, there is provided an apparatus for playing media content by a client in a data communication network. The apparatus includes a controller for sending, to a server, a request for a media content descriptor, receiving, from the server, a media content descriptor including a Uniform Resource Locator (URL) for an access to specific media content data configured for

execution of a specific function, and receiving the specific media content data using the URL included in the received media content descriptor.

### **Brief Description of Drawings**

- [14] FIG. 1 is a diagram illustrating an arrangement of I, P and B frames in conventional single-layer coding;
- [15] FIG. 2 is a diagram illustrating a problematic situation occurring in conventional X-speed play and random access;
- [16] FIG. 3 is a block diagram of a media content playing apparatus according to an embodiment of the present invention;
- [17] FIG. 4 is a flow diagram illustrating an operation between a server and a client in a media content playing apparatus according to an embodiment of the present invention;
- [18] FIGs. 5 to 8 are diagrams illustrating file structures defined by media content descriptors according to an embodiment of the present invention;
- [19] FIGs. 9 is diagrams illustrating a media content playing method according to an embodiment of the present invention;
- [20] FIG. 10 is a flowchart illustrating a media content playing method according to an embodiment of the present invention; and
- [21] FIGs. 11 is diagrams illustrating a media content playing method according to another embodiment of the present invention.

### **Mode for the Invention**

- [22] Embodiments of the present invention will now be described in detail with reference to the accompanying drawings. In the following description, specific details such as detailed configuration and components are merely provided to assist the overall understanding of embodiments of the present invention. Therefore, it should be apparent to those skilled in the art that various changes and modifications of the embodiments described herein can be made without departing from the scope and spirit of the invention. In addition, descriptions of well-known functions and constructions are omitted for clarity and conciseness.
- [23] In the following description, an operation mode in which media content data is played at a normal speed is called a 'normal play mode' or 'normal mode', while an operation mode in which media content data is played at nX-speed (where n is a specific number designating a multiplier of a normal speed, which is equivalent to 1X-speed) upon an X-speed play request of a user is referred to as an 'X-speed play mode' or 'X-speed mode' herein. A play request made by a user to play media content data beginning from a specific time is referred to as a 'random access request', and an operation mode launched by the random access request is referred to as a 'random access mode' herein. A combination of the X-speed play mode and the random access

mode is referred to a 'trick mode' herein.

- [24] Embodiments of the present invention provide a method for minimizing a time delay that may occur in the X-speed mode and/or the random access mode during a media content play operation. In such a method, while media content data received and compressed by a video streaming/download service is being played, if a user requests an X-speed play function (e.g., 2X or 0.5X speed play), or requests to play the media content data beginning from a specific time, a time delay between the user request and a start of the play may be reduced, enabling precise play upon user request.
- [25] In addition, embodiments of the present invention provide a user-friendly operation, when switching from the X-speed play mode and/or random access mode back to the normal play mode.
- [26] MPEG2, MPEG4, H.263, and H.264, which are non-limiting examples of technologies related to embodiments of the present invention, and video compression technologies such as High-efficiency Video Coding (HVC), which have recently emerged as standards for digital media, use the following three schemes to increase compression efficiency:
- [27] A first scheme for increasing compression efficiency includes converting a Red/Green/Blue (RGB) input into a luminance/blue-chrominance/red-chrominance (YCbCr) input. In this scheme, duplicate luminance and chrominance information in the RGB input is converted into Y information and Cb and Cr information, respectively.
- [28] A second scheme for increasing compression efficiency includes removing spatial redundancy of the converted data. In this scheme, the spatial similarity in one screen (or one-screen data) is removed using Discrete Cosine Transform (DCT), quantization, and Variable Length Coding (VLC).
- [29] A third scheme for increasing compression efficiency includes removing temporal redundancy between consecutive screens, based on similarity of parts of successive images that are close to each other in time. This temporal redundancy may be removed by a prediction technique such as Differential Pulse Code Modulation (DPCM), using a Motion Vector (MV) measured by Motion Estimation (ME).
- [30] FIG. 1 illustrates an example of a method for removing temporal redundancy, and shows an arrangement of Intra (I), Predictive (P), Bi-directional (B) frames in general single-layer coding, and reference relationships between them.
- [31] To remove the temporal redundancy, types of frames are set as shown in FIG. 1 in the MPEG technology, and the types of frames are classified into I, P, and B frames. The I frame, an independent frame, can be decoded using only its own data. The P frame is decoded by referencing the preceding I frame. The B frame is decoded by referencing the preceding I frame and the succeeding P frame.

- [32] In every coding scheme, generally, the first screen is coded with an I frame, while the succeeding screens undergo coding with P frames or B frames, in order to reduce a bit rate. However, in the case where only P frames and B frames continuously exist, once an error occurs while the P or B frames are being decoded in a reception side, successive errors may occur. In this case, if a user requests a random access to a specific frame, decoding for a precise random access to the frame may be difficult to perform.
- [33] Therefore, an I frame is periodically inserted between the consecutive P and B frames, which is called 'intra-refresh'. The number of frames between an arbitrary I frame and its next I frame is called a Group of Picture (GOP). In the example of FIG. 1, a GOP value is 9.
- [34] FIG. 2 illustrates possible problems, which may occur when normal-mode data configured for normal play is played in an X-speed play mode or a random access mode.
- [35] Frames on the upper part of FIG. 2 represent normal-mode data configured for normal play. Generally, in the normal-mode data is inserted an I frame for intra-refresh once every 15 frames. In this case, for a video clip configured to include 30 frames per second, i.e., for a video clip configured to have two I frames inserted per second, since a time difference between the two I frames is 0.5 second, if a random access is made on this video clip, a time delay of up to 0.5 second may occur. The same problem may occur not only in the streaming video clips such as broadcasts, but also when video clips having been downloaded and stored in advance, are played.
- [36] Referring to FIG. 2, two instances of accessing normal mode data are considered. More specifically, an instance where random access is performed with normal-mode data, and an instance where X-speed play is performed with normal-mode data.
- [37] Regarding use of a random access mode with normal-mode data, if a random access request is made by a user in a specific frame within a GOP period between an I frame 201 and the next I frame 203, a client starts decoding beginning from the next I frame 203 after moving by a specific time  $d_1$ , instead of starting the decoding beginning from a P frame 205 right after the time 200 the random access request was made. This is because the initial decoding starts at an I frame, as described above. In this case, as the time the random access was requested by the user is different from the time the decoding actually starts, no frames are played in a period between the two times, making it impossible to precisely carry out the random access operation requested by the user.
- [38] Regarding X-speed play with normal-mode data, when the normal-mode data undergoes X-speed play, only I frames are decoded for fast play, making it difficult to efficiently support the X-speed play function for X=non-integer (e.g., 1.2 or 0.4),

rather than for  $X=\text{integer}$  (e.g., 1 and 2).

[39] Therefore, embodiments of the present invention provide a method for supporting an efficient  $X$ -speed play mode, and minimizing time delays occurring when a random access request is made.

[40] FIG. 3 illustrates a block diagram of a media content playing apparatus according to an embodiment of the present invention.

[41] Referring to FIG. 3, a client 350 is connected to an external server 300 over a communication network (not shown). The client 350 is a device for playing, in real time, media content data received from the server 300 by the streaming or download method, or for recording parts of the received media content data in a storage 357, analyzing the stored media content data, and playing the analyzed media content data. The client 350 may include Digital TeleVisions (DTVs), personal computers, Digital Versatile Disc (DVD) players, etc.

[42] Although the client 350 illustrated in FIG. 3 is adapted to store media content data in an embodiment of the present invention, this configuration does not limit the present invention thereto, and the client 350 may have only the function of playing media content data in real time, without the storage function in accordance with embodiments of the present invention. The communication network connecting the client 350 to the server 300 includes any wired/wireless communication networks.

[43] The server 300 may be managed by a broadcasting station, a telecommunication operator, a film producer, or a private content provider. The server 300 may store various complex media content data, such as audio data, video data, text data, image data, etc., and metadata therefore.

[44] Referring to FIG. 3, the client 350 receives media content data from the server 300. However, the client 350 may receive media content data from a database rather than the server 300.

[45] The client 350 includes a communication unit 351, the storage 357, an analyzer 355, a decoder 361, a controller 353, and an output unit 359. Although not illustrated in FIG. 3, the client 350 may further include a user interface responsible for interfacing with a user. The user may send a request to the controller 353 using the user interface. Generally, the user interface, which receives user inputs, may include a physical conversion unit, such as a keyboard, a mouse, a touch screen, a haptic device, and a microphone. The user interface receives command, text, event, number, or voice information from the user, and provides the received information to the controller 353.

[46] According to embodiments of the present invention, the user may control the controller 353 using the user interface to play media content data corresponding to the current GOP period in an  $X$ -speed play mode, and may also set a play mode in which the user desires to play media content data of the next GOP period, while the media

content data of the current GOP period is being played in the X-speed play mode.

- [47] The communication unit 351 is connected to the server 300 or a third database (not shown) over a wired/wireless network, and exchange data with the server 300 or the third database. If the client 350 uses wireless network technology, the communication unit 351 may transmit/receive control commands and data between the server 300 and the client 350 using the common data packet communication method. The communication method may include mobile communication technologies, such as Wireless Local Area Network (WLAN), Bluetooth, Zigbee, Wireless Broadband (WiBro), Wireless Media (WiMedia), 3<sup>rd</sup> Generation (3G) Code Division Multiple Access (CDMA).
- [48] The storage 357 stores the information that the communication unit 351 has received from the server 300 or the third database. The analyzer 355 analyzes media content data stored in the storage 357, and provides the analyzed media content data to the decoder 361.
- [49] The decoder 361 decodes the media content data analyzed by the analyzer 355 using an inverse method of the coding method based on an MPEG file format according to embodiments of the present invention, and provides the decoded media content data to the output unit 359. If the media content data includes different types of data, the decoder 361 selects required media content data and provides it to the output unit 359.
- [50] For example, if the media content data includes 'normal-mode data' for the normal play mode and 'trick-mode data' for the X-speed play mode or the random access mode, and the X-speed play mode is selected, then the controller 353 may select the trick-mode data and provide the trick-mode data to the decoder 361.
- [51] The output unit 359 may output the media content data decoded by the decoder 361 to the screen. The output unit 359 may output the overall condition of the client 350 and the information that the user has input through the user interface, to the screen or the speaker.
- [52] Although not illustrated in FIG. 3, the client 350 operates based on system time information.
- [53] While media content data of the initial GOP period is played in an X-speed play mode, the user may set a play mode for playing media content data of the next GOP period, using the user interface. In some cases, the client 350 may decode media content data of the next GOP period using a method predefined therein, regardless of user operations.
- [54] Table 1 shows a structure of a media content descriptor according to an embodiment of the present invention. The media content descriptor is information provided from a server or the third database to a terminal or a client, and includes various information that the terminal (or client) requires to perform a function associated with its media

content data.

[55] For example, referring to Table 1, a media content descriptor according to an embodiment of the present invention includes not only information for the normal play mode, but also information for the X-speed play mode and the random access mode.

[56] The media content descriptor may be included in media content data when the media content service is initially launched, or included in the media content data periodically, and then transmitted from the server or the third database to the terminal (or client). In one example, only part of the information included in the media content descriptor may be transmitted to the terminal. In this case, the terminal may perform only the function corresponding to the transmitted information.

[57] Table 1

[Table 1]

```

<-----mediaContentsManifest.xml-----><Contents
xsi:noNamespaceSchemaLocation="Samsung_mmt.xsd"xmlns:xsi=
http://www.w3c.org/2001/XMLSchema-instance> <ContentItem> <Title
xml:lang="en-us">English</Title> <Synopsis xml:lang="en-us">String</Synopsis>
<OriginSite> http://samsungmmtexample.com</OriginSite>
<OriginSiteName>SamsungExample<OriginSiteName>
<ContentID>123123109</ContentID> <ContentURL VideoCoding="AVC"
Duration="14:30:00:00" MediaFormat="MP4"DRMSystemID="128"
MD5Hash="String" MIMEType="video/mp4"
TransferType="AS"AudioCoding="AAC"> <ContentURL>
http://samsungexample.com/vod/movies/123123109/Description/
MainContents/
MainContentsDescription.xml </ContentURL>
<-----mediaFileStructure-----> <ContentType>
<URLTemplate></URLTemplate> <NextControlURL></NextControlURL>
<RelDataURL></RelDataURL> <Track ID=1 GOP="9" StartTime=0.00
mode=normal Bitrate=40000 SegmentStartID=01SegmentDuration=1 Seg-
menCount=1> <Segment ID=1 StartTime=0.00 Duration=1></SegmentID> <Track
ID=2 GOP="3" StartTime=0.00 mode=Trickplay Bitrate=10000 Seg-
mentStartID=01SegmentDuration=1 SegmenCount=1> <Segment ID=1
StartTime=0.00 Duration=1></SegmentID> </Track> </ContentType>

```

[58] The media content descriptor is described in further detail as follows with reference to Table 1.

[59] The media content descriptor includes manifest information and file structure information.

[60] As an element of a 'Contents' attribute representing Schema information being

referenced by the media content descriptor, the manifest information of the media content includes at least one of 'Title' information, 'Synopsis' information, 'OriginSite' information including creator information of the media content, 'OriginSiteName' containing a name of the creator of the media content, 'ContentID' for identifying the media content, 'ContentsURL' information by which control information of the media content may be referenced, and 'ContentItem' information representing the media content itself.

- [61] The file structure information, which represents a file structure of the media content data, includes as a container element of a 'ContentType' attribute, information such as 'URLTemplate' representing a positioning method of the actual media content data in the streaming or download method, 'NextControlURL' which is controlled based on the next GOP unit, 'RelDataURL' by which additional information such as X-speed information and information about the time supportable in random access, can be referenced, and 'TrackID' by which normal-mode data available in the normal mode and trick-mode data available in the X-speed play mode or the random access mode can be identified.
- [62] The controller 353 of the client 350 may analyze a MainContentsDescription.xml file with ContentURL included in the manifest information, and send a request for desired media content data information to the server 300 on a GOP basis. For example, while viewing a video clip in the normal play mode at an initial state, a user may switch from the normal play mode to the X-speed play mode or the random access mode upon requesting the X-speed play function or the random access function. In this case, the client 350 sends, to the server 300, a request for manifest information for the X-speed play or the random access. Upon receiving the manifest information in Table 1 from the server 300, the client 350 may send a request for a MainContentsDescription.xml file to an address included in the ContentsURL information included in the manifest information, and receive trick-mode data for the X-speed play or random access from the address.
- [63] FIG. 4 illustrates an operation between a server and a client with a media content playing apparatus according to an embodiment of the present invention.
- [64] Referring to FIG. 4, the operation is performed between the server 300 and the client 350 in accordance with the data file structure described in Table 1. According to process, if a user of a client desires the X-speed play or random access while receiving specific media content data from a server and playing the received data, the client sends a request for manifest information to the server, receives the manifest information from the server in response to the request, analyzes the received information, and requests the X-speed play or random access function if the analyzed information indicates that the server supports the X-speed play or random access function. It is

assumed in FIG. 4 that the X-speed play is requested.

[65] More specifically, if a user of the client 350 desires X-speed play while receiving and playing specific media content data, the client 350 requests the server 300 to transmit manifest information in step 401. In response, the server 300 transmits the requested manifest information to the client 350 in step 403.

[66] In step 405, the client 350 analyzes a MainContentsDescription.xml file with ContentURL included in the received manifest information, and sends a request for the X-speed play function to the address (e.g., a URL) of the server 300, included in the ContentURL, if the analyzed information indicates that the server supports the X-speed play function. In response, the server 300 transmits a fragment box related to the X-speed play to the client 350 in step 407. For reference, the term 'fragment box' refers to collection of information required for execution of a specific function. The client 350 performs X-speed play using information included in the fragment box.

[67] FIGs. 5 to 8 illustrate configurations of normal-mode data and trick-mode data according to an embodiment of the present invention.

[68] As previously stated herein, the normal-mode data is data available in the normal play mode, while the trick-mode data is data available in the X-speed play mode or the random access mode. According to embodiments of the present invention, the trick-mode data is generally configured to be  $1/4$  times or  $1/N$  (where  $N$  is several tens, i.e., 10, 20, 30, etc.) times the normal-mode data in size, and the GOP value of the trick-mode data is also set smaller than that the GOP value of the normal-mode data, increasing the frequency of I frames (or the number of occurrences of I frames). In other words, the trick-mode data has an overall lower amount of data than the normal-mode data, while I frames of the trick-mode data appear more frequently than I frames of normal-mode data. Therefore, use of the trick-mode data increases the speed of transmission/reception and decoding, compared with use of normal-mode data. As a result, the time delay in the X-speed play and random access is decreased, and enables a user to obtain precise playback and/or search of media data. According to embodiments of the present invention, the normal-mode data may still be played in either of the X-speed play mode or the random access mode, while the trick-mode data may be played in the normal play mode. However, it is assumed herein that the user will elect to use the normal-mode data in the normal play mode, and the trick-mode data in the X-speed play mode or the random access mode.

[69] FIG. 5(a) illustrates normal-mode data with  $GOP=9$ , and FIG. 5(b) illustrates trick-mode data with  $GOP=3$ . The trick-mode data in FIG. 5(b) is smaller than the normal-mode data in FIG. 5(a) in terms of both the GOP value and the data size.

[70] FIGs. 6 to 8 illustrate other configurations of normal-mode data and trick-mode data according to an embodiment of the present invention.

- [71] FIGs. 6(a) to 6(c) illustrate an example, where the data in FIGs. 6(a) and 6(b) is identical to the data in FIGs. 5(a) and 5(b), respectively. FIG. 6(c) illustrates trick-mode data configured to include only I frames. If the trick-mode data includes only I frames, the play speed in the X-speed play and the shift speed in the random access may be further increased.
- [72] FIGs. 7(a) to 7(c) illustrate another example, where FIG. 7(a) illustrates an example of complex (or multi-layer) normal-mode data. For example, if each frame includes 3-layer data, data scalability may be achieved using the 3-layer data. Data in FIGs. 7(b) and 7(c) is equal to the data in FIGs. 6(b) and 6(c), respectively.
- [73] FIGs. 8(a) to 8(c) illustrate yet another example, in which trick-mode data is realized in scalable data. That is, normal-mode data in FIG. 8(a) is equal to the normal-mode data in FIG. 5(a). However, the complex (or multi-layer) trick-mode data in FIGs. 8(b) and 8(c) are configured by combining multiple sets of the trick-mode data in FIGs. 6(b) and 6(c), respectively, to support the scalability thereof.
- [74] As illustrated in FIGs. 7 and 8, the media content data of the present invention may include multi-layer normal-mode data (FIG. 7(a)) and/or multi-layer trick-mode data (FIGs. 8(b) and 8(c)). When configured in a multi-layer data structure, the media content data may support scalability. As the normal-mode data and the trick-mode data are presented independently from each other, they may be played in different times independently.
- [75] As described above, the user plays the normal-mode data in the normal mode. Also, the user may choose to play the trick-mode data in the X-speed play mode or the random access mode. For example, while the normal-mode data is being played in the normal mode, if the user chooses the X-speed play mode, the client requests trick-mode data and plays the trick-mode data in the X-speed play mode.
- [76] While the normal-mode data is being played in the normal mode, if the user chooses to play the trick-mode data in the X-speed play mode, the controller 353 analyzes system time clock information, by which the controller 353 provides time information in the normal-mode data to the output unit 359 and the decoder 361, and then requests the trick-mode data according to the analysis results. Therefore, if X2-speed play is requested during play of the normal-mode data, the client plays the trick-mode data at a double speed.
- [77] FIGs. 9(a) and 9(b) illustrate an operation of performing random access using trick-mode data according to an embodiment of the present invention. FIG. 9(a) illustrates normal-mode data, while FIG. 9(b) illustrates trick-mode data.
- [78] Referring to FIGs. 9(a) and 9(b), while the normal-mode data (as shown in FIG. 9(a)) is being played, a random access request is made by a user at a time 901. Conventionally, when random access is made on the normal-mode data (FIG. 9(a)), decoding

starts beginning from the next I frame 903 rather than the P frame 907 succeeding the time 901, causing a time delay of  $d_1$ . However, when random access is made on the trick-mode data (as shown in FIG. 9(b)), decoding may start beginning from the I frame 905 just next to the time 901, reducing the time delay to  $d_2$ , which is remarkably shorter than the conventional delay time  $d_1$ .

[79] FIG. 10 illustrates a media content playing method according to an embodiment of the present invention, in which a client 350 has sent a request for a media content descriptor to a server 300 to play normal-mode data.

[80] Referring to FIG. 10, in step 701, the client 350 analyzes a media content descriptor that the client 350 has received from the server 300 in response to the request. In step 703, if information about normal-mode data is included in the media content descriptor, the client 350 requests/receives normal-mode data from the server 300, and decodes the received normal-mode data.

[81] In step 705, the client 350 determines whether a switching request to an X-speed play mode or a random access mode has been received from a user while the normal-mode data is being played. Upon receiving the switching request, the client 350 stops decoding the normal-mode data and proceeds to step 707.

[82] In step 707, the client 350 sends a request for trick-mode data to a URL according to track information (i.e., a server's URL address included in 'ContentURL' described herein with reference to FIG. 4) included in the media content descriptor, and receives the requested trick-mode data. In step 709, the client 350 decodes media content data corresponding to the received trick-mode data.

[83] However, upon failure to receive the switching request in step 705, the client 350 proceeds directly to step 709, and continues to decode and play the normal-mode data in the normal play mode.

[84] FIGs. 11(a) and 11(b) illustrate a media content playing method according to another embodiment of the present invention.

[85] In the example of FIGs. 11(a) and 11(b), a user has chosen not to play normal-mode data in an X-speed play mode or a random access mode. While trick-mode data is played in the X-speed play mode or random access mode, if the user requests switching from the X-speed play mode or random access mode to the normal mode, the client performs mode switching and plays the trick-mode data in the normal mode according to an embodiment of the present invention. FIG. 11(a) illustrates normal-mode data, while FIG. 11(b) illustrates trick-mode data.

[86] As assumed above, the client 350 is playing media content using the trick-mode data (as shown in FIG. 11(b)) in the X-speed play mode or random access mode. While the trick-mode data (as shown in FIG. 11(b)) is played, if the user requests switching to the normal mode at a time 1001, the client 350 stops playback of the trick-mode data (i.e.,

trick play) and playback the normal-mode data (FIG. 11(a)) at the time 1001. In this case, however, because decoding of the normal-mode data starts beginning from an I frame as described above, a time difference occurs between an end time 1001 of the trick play and a first I frame 1003 in the normal-mode data after the end time 1001 of the trick play, resulting in a delay time. No content is played during this delay time.

[87] According to another embodiment of the present invention, to ensure continuous content play for the delay time, the existing trick-mode data is used to continue playback during this delay time. More specifically, during the time period between the end time 1001 of the trick play and the first I frame 1003 in the normal-mode data, frames 1005 of the trick-mode data are decoded in order of I==>P==>P frames. Thereafter, if reception of the first I frame 1003 in the new GOP period in the normal-mode data is completed, the normal-mode data is decoded to play content.

[88] According to another embodiment of the present invention, the media content playing apparatus described with reference to FIG. 3 operates as follows.

[89] The controller 353 determines whether a switching request from an X-speed play mode or a random access mode to the normal mode has been received from a user. Upon receiving the switching request to the normal mode, the controller 353 informs the analyzer 355 of the switching request. The analyzer 355 analyzes ContentURL of the server 300, included in manifest information of a received media content descriptor, and provides a URL of the server 300 to the controller 353 according to the analysis results. The controller 353 sends a request for normal-mode data to the URL on a GOP basis beginning from a new I frame.

[90] The decoder 361 continues to decode the existing trick-mode data until the decoder 361 receives I-frame data in a new GOP period of the normal-mode data from the server 300 and decodes the received frame data. Upon receiving a first I frame in the new GOP period of the normal-mode data, the controller 353 stops the decoding of the trick-mode data, and controls the decoder 361 to decode the normal-mode data. The decoder 361 decodes the normal-mode data under control of the controller 353.

[91] As the manifest information described in the present invention is used to provide information for provision of a media content service, service setting, and/or service initialization, the manifest information may be alternatively be referred to as 'configuration information', for example. However, any other names are also available as long as information thereof has the same meanings or roles as the manifest information. Although basic fields and/or their attribute information needed to describe embodiments of the present invention have been described so far, information of a configuration file described in each embodiment of the present invention may further include additional information as fields and attributes according to the environments or services. Also, while logical means of information included in the configuration file

have been illustrated and described so far, the information may be represented in the form of various information representation structures or formats, such as eXtensible Markup Language (xml), binary data, and a formatted information structure, according to the system and service production/provision environment.

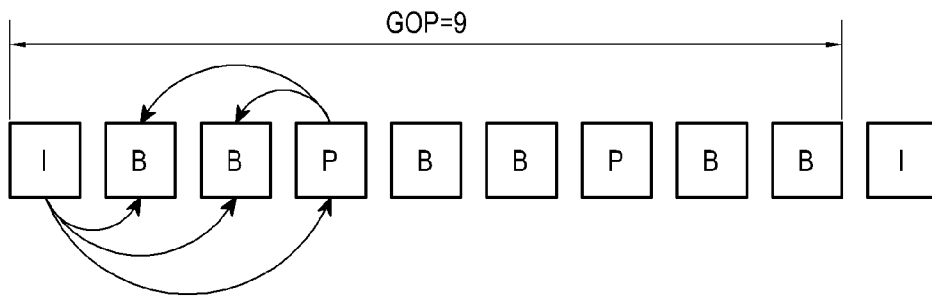
- [92] As is apparent from the foregoing description, in providing complex content including multiple source contents or multiple streams, the present invention sets a repeat cycle of intra-refresh frames for trick-mode data to be shorter than a repeat cycle of intra-refresh frames for normal-mode data in a video service such as Digital TV that uses the streaming and download methods, thereby supporting a precise X-speed play function and minimizing an increase in bit rate and a time delay during random access.
- [93] In addition, according to embodiments of the present invention, by setting time points of intra-refresh frames in a trick-mode data layer that are different from time points of intra-refresh frames in a basic layer, the bit rate is prevented from being concentrated on intra-frames and forming a peak, facilitating planarization of the bit rate.
- [94] Another effect of embodiments of the present invention reduces complexity of Picture-In-Picture (PIP) screens to  $1/4$  or  $1/N$  (where  $N$  is several tens), PIP may be simply implemented, and if the available bit rate is variable like in the wireless network or Internet Protocol (IP) network, only separate trick-mode data may be transmitted according to the network situation.
- [95] Meanwhile, various other effects of the present invention have been disclosed explicitly or implicitly in the foregoing detailed description of embodiments of the present invention.
- [96] While the invention has been shown and described with reference to certain embodiments thereof, it will be understood by those skilled in the art that various changes in form and details may be made therein without departing from the spirit and scope of the invention as defined by the appended claims and their equivalents.

## Claims

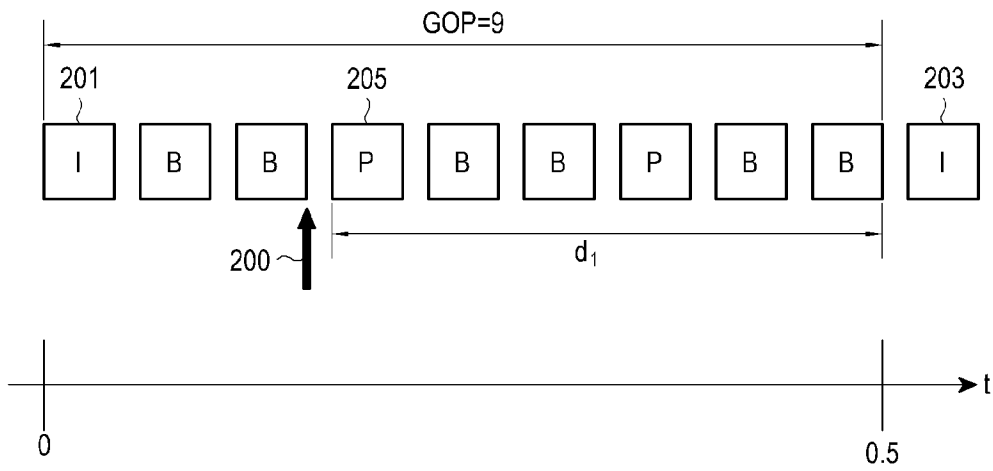
- [Claim 1] A method for playing media content by a client in a data communication network, comprising: sending, to a server, a request for a media content descriptor; receiving, from the server, a media content descriptor including a Uniform Resource Locator (URL) for access to specific media content data configured for execution of a specific function; and receiving the specific media content data using the URL included in the received media content descriptor.
- [Claim 2] The method of claim 1, wherein the received media content descriptor includes a Group Of Picture (GOP) or period information of the specific media content data.
- [Claim 3] The method of claim 1, wherein the specific media content data configured for execution of the specific function includes trick-mode data configured for execution of an X-speed play function or a random access function.
- [Claim 4] The method of claim 3, wherein the trick-mode data includes data having a quality that is lower than a quality of normal-mode data configured for execution of a normal play mode.
- [Claim 5] The method of claim 1, wherein the receiving the specific media content data using the URL comprises: sending a request for the specific media content data to an address of the URL; receiving the specific media content data from a server corresponding to the address of the URL; and performing the specific function using the received specific media content data.
- [Claim 6] A method for supporting media content play of a client, by a server in a data communication network, comprising: receiving, from the client, a request for a media content descriptor; and transmitting, to the client, a media content descriptor including a Uniform Resource Locator (URL) for access to specific media content data configured for execution of a specific function.
- [Claim 7] The method of claim 6, wherein the transmitted media content descriptor includes a Group Of Picture (GOP) or period information of the specific media content data.
- [Claim 8] The method of claim 6, wherein the specific media content data configured for execution of the specific function comprises trick-mode data configured for execution of an X-speed play function or a random access function.

- [Claim 9] The method of claim 8, wherein the trick-mode data includes data having lower quality than a quality of normal-mode data configured for execution of a normal play mode.
- [Claim 10] The method of claim 6, further comprising:receiving from the client a request for the specific media content data associated with an address of the URL; andtransmitting the specific media content data stored according to the address of the URL, to the client.
- [Claim 11] An apparatus for playing media content by a client in a data communication network, comprising:a controller for sending, to a server, a request for a media content descriptor, receiving, from the server, a media content descriptor including a Uniform Resource Locator (URL) for access to specific media content data configured for execution of a specific function, and receiving the specific media content data using the URL included in the received media content descriptor.
- [Claim 12] The apparatus of claim 11, wherein the received media content descriptor includes a Group Of Picture (GOP) or period information of the specific media content data.
- [Claim 13] The apparatus of claim 11, wherein the specific media content data configured for execution of the specific function includes trick-mode data configured for execution of an X-speed play function or a random access function.
- [Claim 14] The apparatus of claim 13, wherein the trick-mode data includes data having a lower quality than a quality of normal-mode data configured for execution of a normal play mode.
- [Claim 15] The apparatus of claim 11, wherein the controller sends a request for the specific media content data to an address of the URL, receives the specific media content data from a server having the address of the URL, and performs the specific function using the received specific media content data.

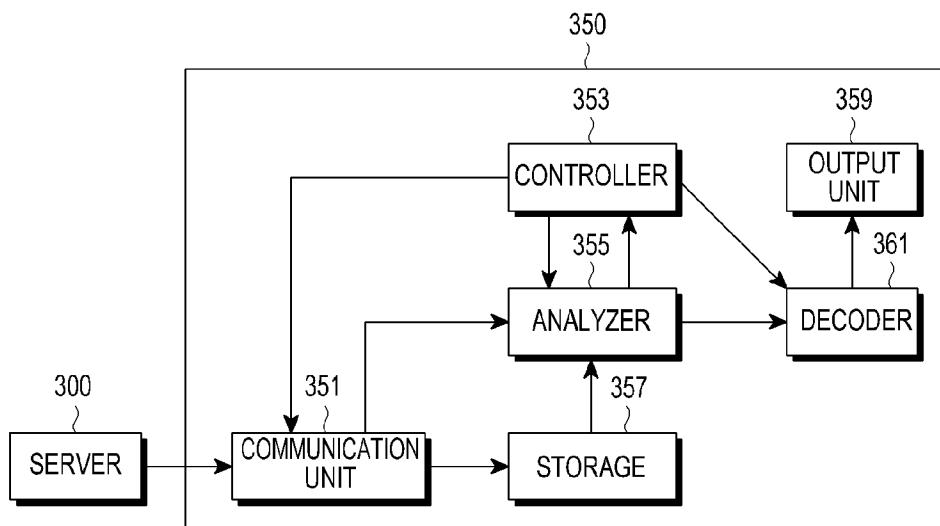
[Fig. 1]



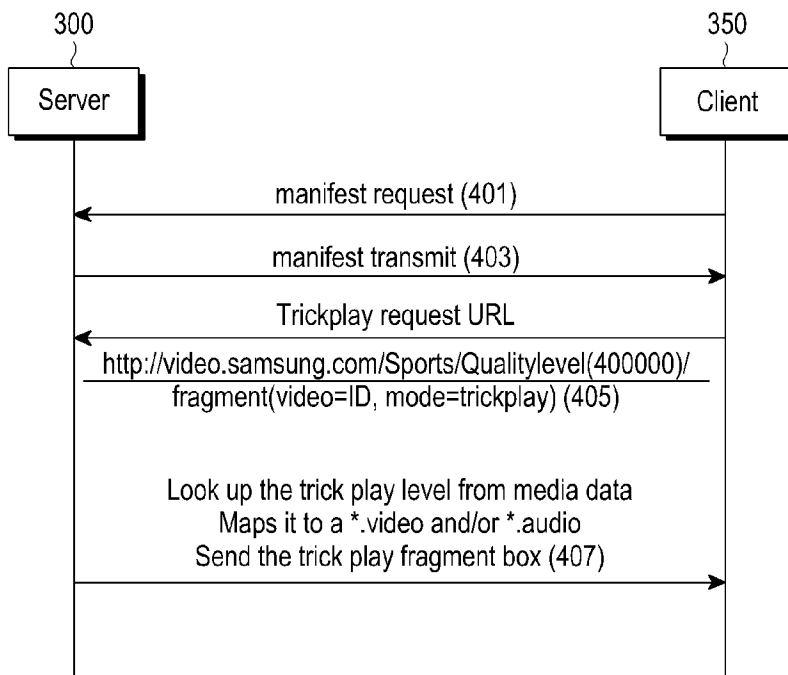
[Fig. 2]



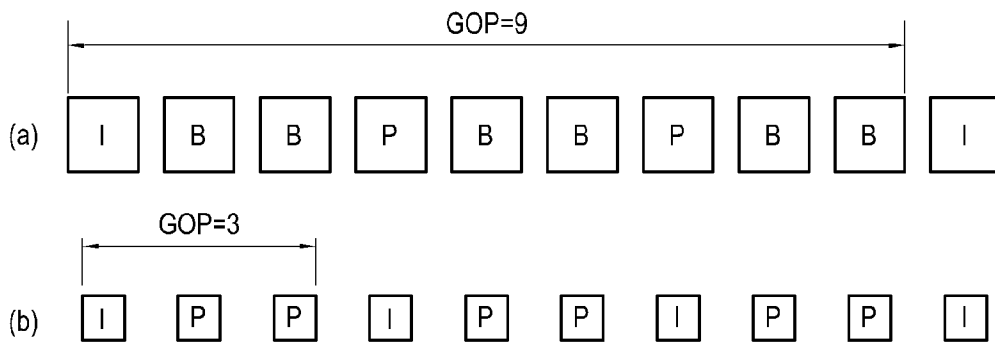
[Fig. 3]



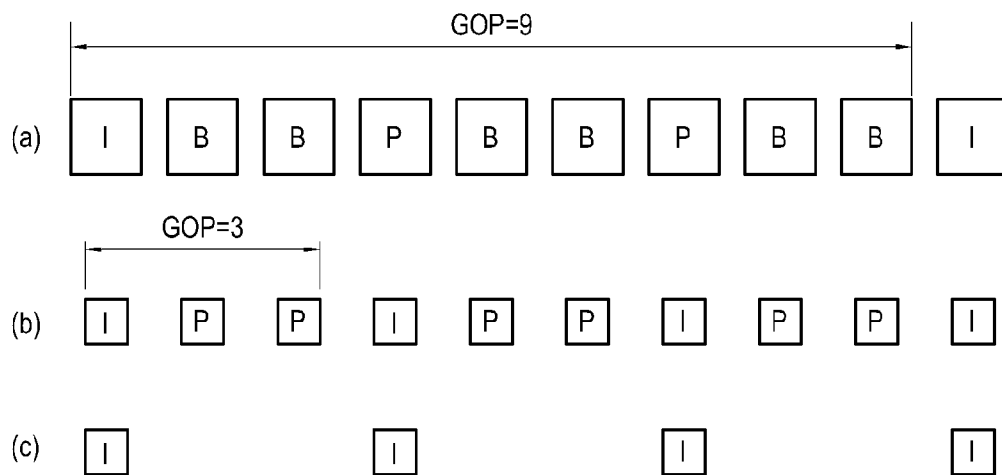
[Fig. 4]



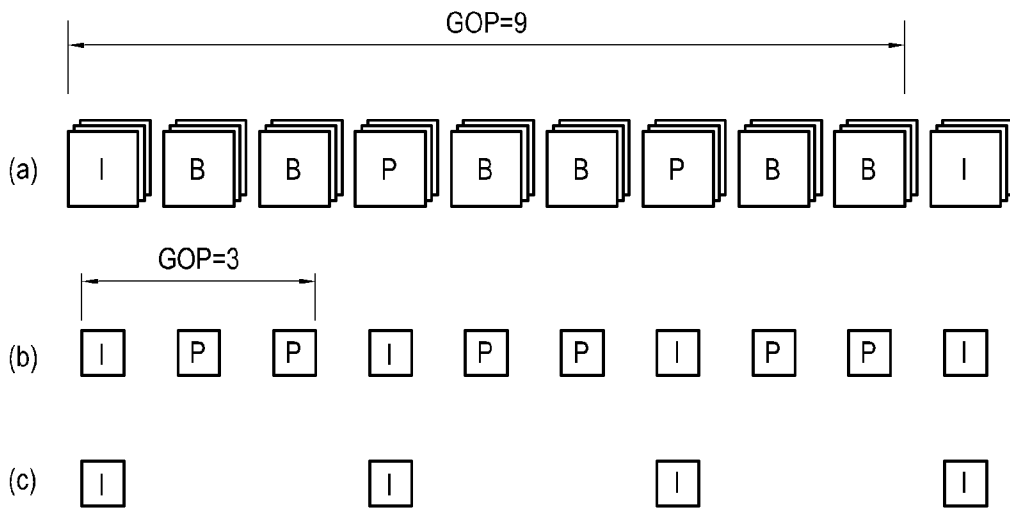
[Fig. 5]



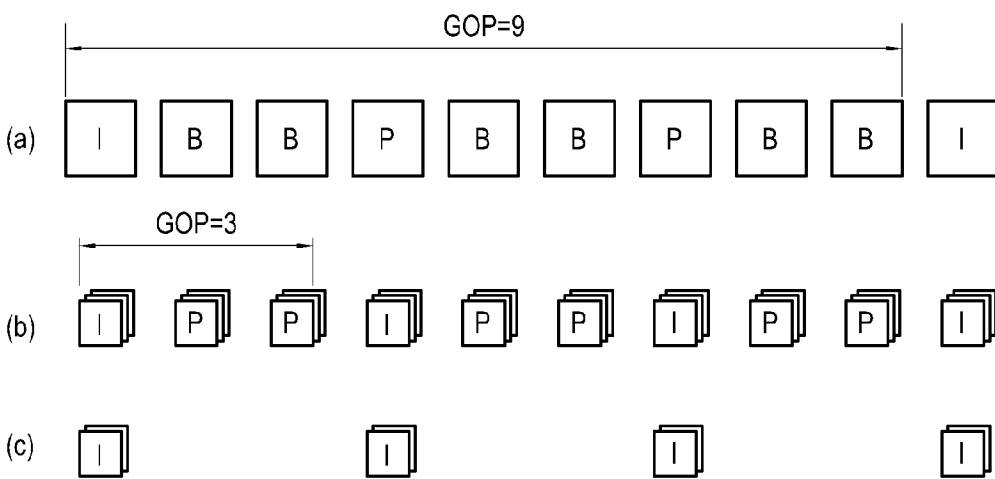
[Fig. 6]



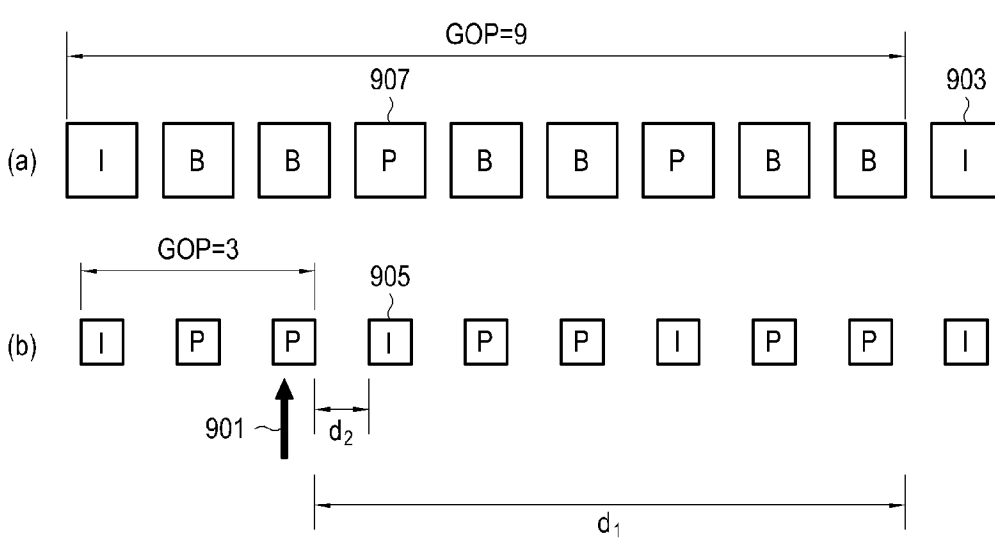
[Fig. 7]



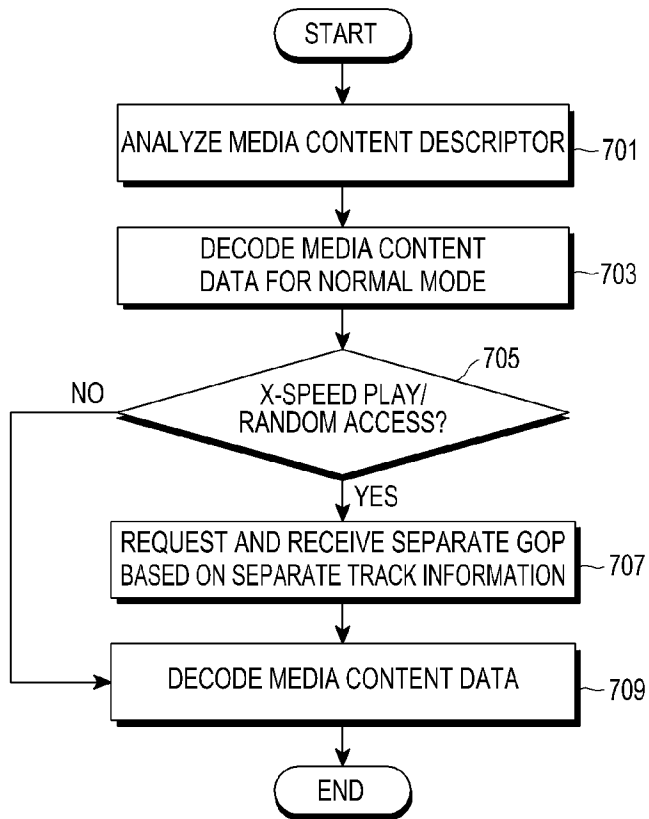
[Fig. 8]



[Fig. 9]



[Fig. 10]



[Fig. 11]

