

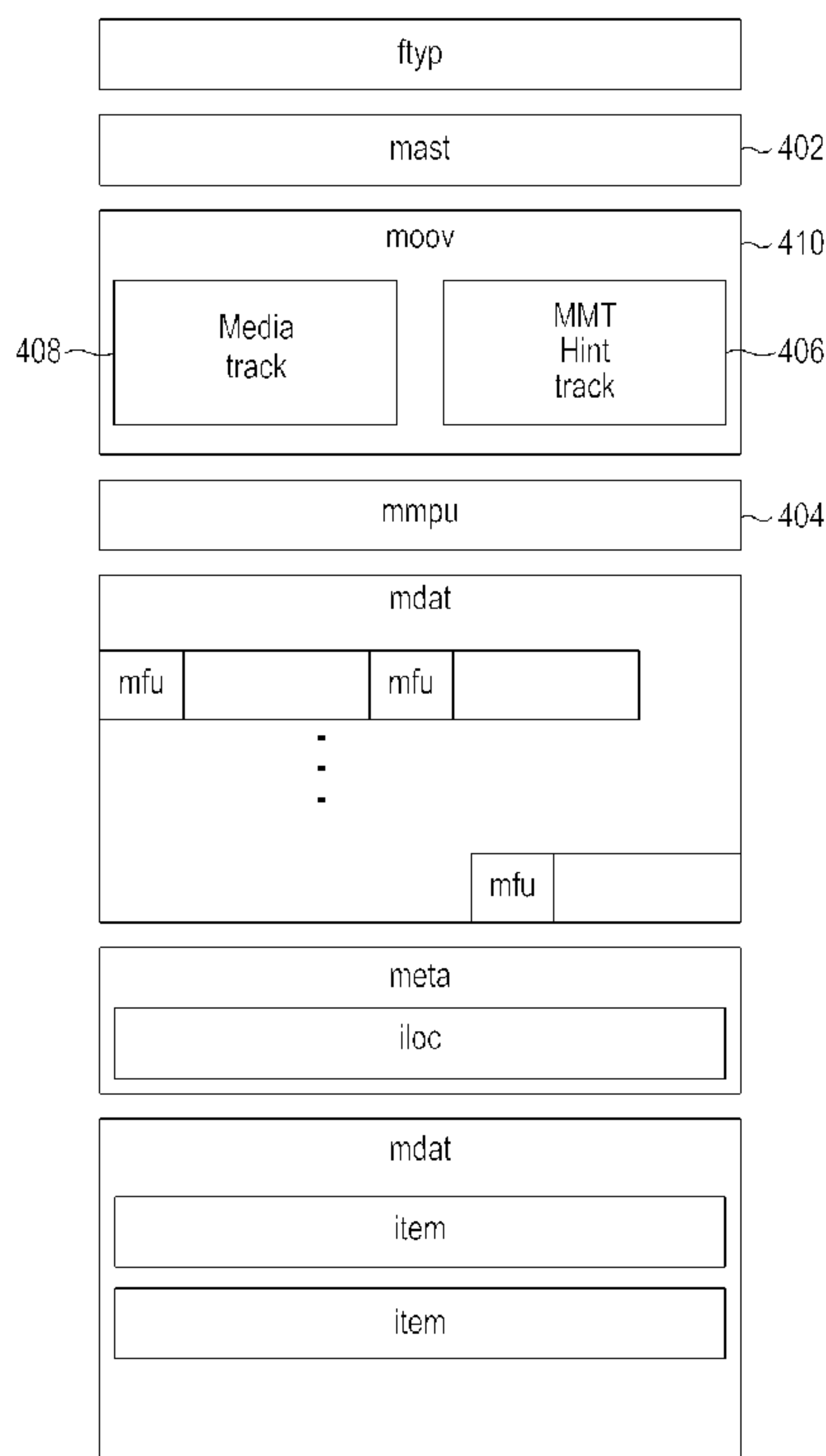


(86) Date de dépôt PCT/PCT Filing Date: 2013/11/28
 (87) Date publication PCT/PCT Publication Date: 2014/06/05
 (85) Entrée phase nationale/National Entry: 2015/05/29
 (86) N° demande PCT/PCT Application No.: KR 2013/010932
 (87) N° publication PCT/PCT Publication No.: 2014/084643
 (30) Priorités/Priorities: 2012/11/29 (US61/731,360);
 2013/11/26 (US14/091,243)

(51) Cl.Int./Int.Cl. *H04N 21/235* (2011.01),
H04N 7/24 (2011.01)
 (71) Demandeur/Applicant:
 SAMSUNG ELECTRONICS CO., LTD., KR
 (72) Inventeurs/Inventors:
 BOUAZIZI, IMED, US;
 LIM, YOUNG-KWON, US;
 BHAT, KONG POSH, US
 (74) Agent: MARKS & CLERK

(54) Titre : PROCÉDE ET APPAREIL POUR L'ENCAPSULATION D'ACTIFS DE TRANSPORT MULTIMEDIA DE GROUPE D'EXPERTS POUR LE CODAGE D'IMAGES ANIMEES DANS UNE ORGANISATION INTERNATIONALE POUR DES FICHIERS MULTIMEDIAS BASES SUR LA NORMALISATION

(54) Title: METHOD AND APPARATUS FOR ENCAPSULATION OF MOTION PICTURE EXPERTS GROUP MEDIA TRANSPORT ASSETS IN INTERNATIONAL ORGANIZATION FOR STANDARDIZATION BASE MEDIA FILES



(57) **Abrégé/Abstract:**

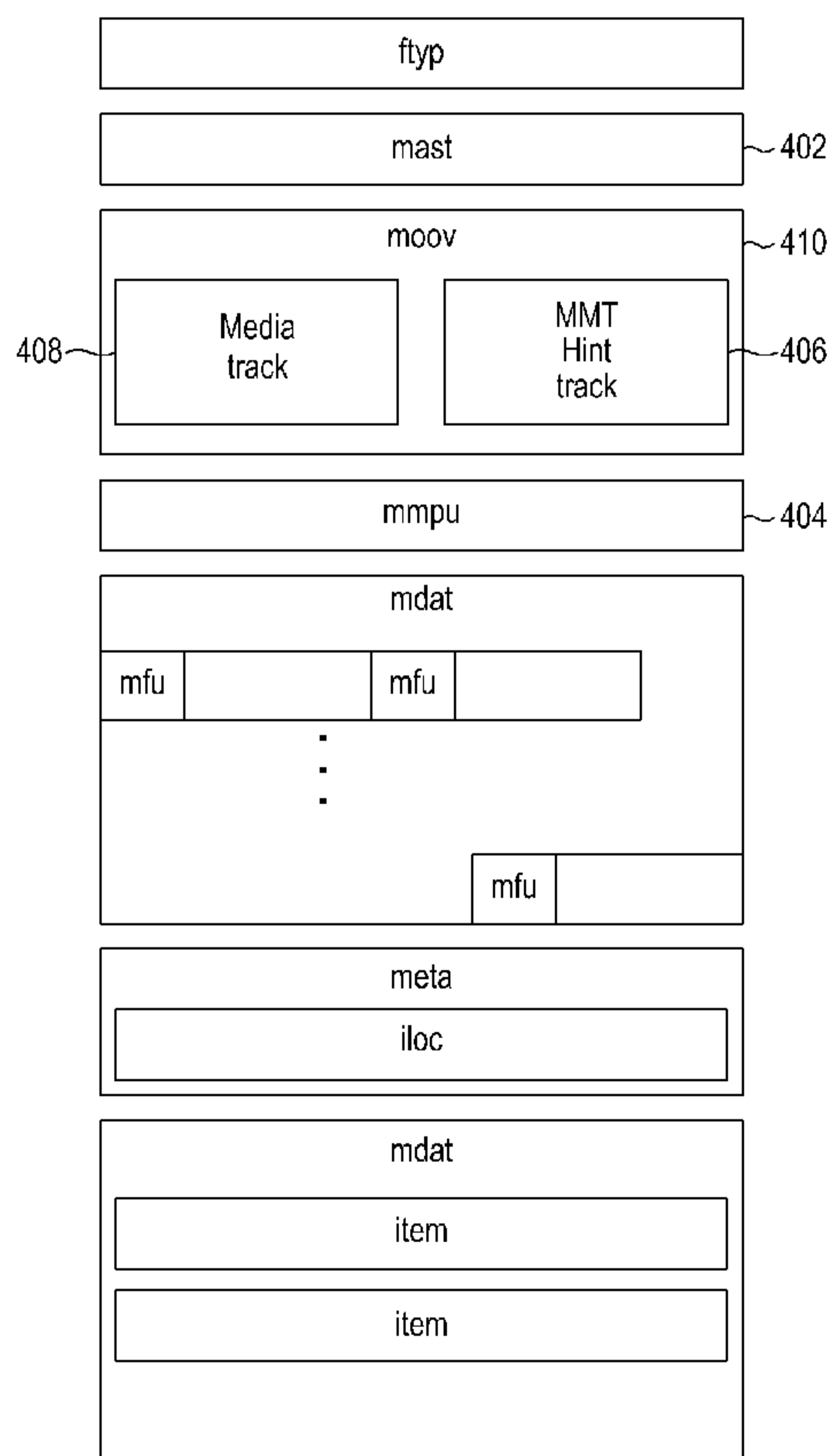
An apparatus includes receive path circuitry configured to receive a Motion Picture Experts Group (MPEG) Media Transport (MMT) container and a processing device configured to identify locations of one or more media fragment units (MFUs) in the MMT container using a hint track within the MMT container. Another apparatus includes transmit path circuitry configured to transmit an MMT container and a processing device configured to identify locations of one or more MFUs in the MMT container using a hint track within the MMT container.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property
Organization
International Bureau(43) International Publication Date
5 June 2014 (05.06.2014)(10) International Publication Number
WO 2014/084643 A1

- (51) **International Patent Classification:**
H04N 21/235 (2011.01) *H04N 7/24* (2011.01)
- (21) **International Application Number:**
PCT/KR2013/010932
- (22) **International Filing Date:**
28 November 2013 (28.11.2013)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
61/731,360 29 November 2012 (29.11.2012) US
14/091,243 26 November 2013 (26.11.2013) US
- (71) **Applicant: SAMSUNG ELECTRONICS CO., LTD.**
[KR/KR]; 129, Samsung-ro, Yeongtong-gu, Suwon-si,
Gyeonggi-do 443-742 (KR).
- (72) **Inventors: BOUAZIZI, Imed;** 8109 Haning Drive, Plano,
Collin County, Texas 75025 (US). **LIM, Young-Kwon;**
9607 Custer Road, Apt. 121, Plano, Collin County, Texas
75025 (US). **BHAT, Kong Posh;** 7405 Stoney Point
Drive, Plano, Collin County, Texas 75025 (US).
- (74) **Agents: LEE, Keon-Joo et al.;** Mihwa Bldg. 110-2, My-
ongryun-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, BN, 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, IR, IS, JP, KE, KG, KN, KP, KR,
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, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA,
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,

[Continued on next page]

(54) **Title:** METHOD AND APPARATUS FOR ENCAPSULATION OF MOTION PICTURE EXPERTS GROUP MEDIA
TRANSPORT ASSETS IN INTERNATIONAL ORGANIZATION FOR STANDARDIZATION BASE MEDIA FILES

(57) **Abstract:** An apparatus includes receive path circuitry configured to receive a Motion Picture Experts Group (MPEG) Media Transport (MMT) container and a processing device configured to identify locations of one or more media fragment units (MFUs) in the MMT container using a hint track within the MMT container. Another apparatus includes transmit path circuitry configured to transmit an MMT container and a processing device configured to identify locations of one or more MFUs in the MMT container using a hint track within the MMT container.

WO 2014/084643 A1



GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, 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, KM, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

Description

Title of Invention: METHOD AND APPARATUS FOR ENCAPSULATION OF MOTION PICTURE EXPERTS GROUP MEDIA TRANSPORT ASSETS IN INTERNATIONAL ORGANIZATION FOR STANDARDIZATION BASE MEDIA FILES

Technical Field

- [1] This disclosure relates generally to a method and apparatus transferring media content for encapsulation of Motion Picture Experts Group (MPEG) Media Transport (MMT) assets in International Organization for Standardization (ISO) Base Media Files.

Background Art

- [2] Motion Picture Experts Group (MPEG) Media Transport (MMT) specifies technologies for the delivery of coded media data for multimedia services over heterogeneous Internet Protocol (IP) network environments. The delivered coded media data includes both (i) audio-visual media data of which the data units have an inherent timeline (namely timed data) and (ii) other types of data that do not have an inherent timeline among their data units (namely non-timed data).

Disclosure of Invention

Technical Problem

- [3] MMT is designed under the assumption that the coded media data is delivered through a packet-based delivery network using Internet Protocols, such as the Realtime Transport Protocol (RTP), Transmission Control Protocol (TCP), and User Datagram Protocol (UDP).

Solution to Problem

- [4] This disclosure provides a method and apparatus for encapsulation of Motion Picture Experts Group (MPEG) Media Transport (MMT) assets in International Organization for Standardization (ISO) Base Media Files.
- [5] In a first embodiment, an apparatus includes receive path circuitry configured to receive an MMT container. The apparatus also includes a processing device configured to identify locations of one or more media fragment units (MFUs) in the MMT container using a hint track within the MMT container.
- [6] In a second embodiment, an apparatus includes transmit path circuitry configured to transmit an MMT container. The apparatus also includes a processing device configured to identify locations of one or more media fragment units (MFUs) in the MMT container using a hint track within the MMT container.
- [7] In a third embodiment, a method includes receiving an MMT container. The method

also includes identifying locations of one or more media fragment units (MFUs) in the MMT container using a hint track within the MMT container.

[8] In a fourth embodiment, a method includes receiving MFUs of an MMT container, identifying missing MFUs, marking them appropriately or adjusting the metadata for correct processing.

[9] Other technical features may be readily apparent to one skilled in the art from the following figures, descriptions, and claims.

[10] Before undertaking the DETAILED DESCRIPTION below, it may be advantageous to set forth definitions of certain words and phrases used throughout this patent document. The terms "include" and "comprise," as well as derivatives thereof, mean inclusion without limitation. The term "or" is inclusive, meaning and/or. The phrase "associated with," as well as derivatives thereof, may mean to include, be included within, interconnect with, contain, be contained within, connect to or with, couple to or with, be communicable with, cooperate with, interleave, juxtapose, be proximate to, be bound to or with, have, have a property of, have a relationship to or with, or the like. The term "controller" means any device, system or part thereof that controls at least one operation, and such a device may be implemented in hardware or in a combination of hardware and software/firmware. It should be noted that the functionality associated with any particular controller may be centralized or distributed, whether locally or remotely. The phrase "at least one of," when used with a list of items, means that different combinations of one or more of the listed items may be used, and only one item in the list may be needed. For example, "at least one of: A, B, and C" includes any of the following combinations: A, B, C, A and B, A and C, B and C, and A and B and C. Definitions for certain words and phrases are provided throughout this patent document, and those of ordinary skill in the art should understand that in many if not most instances, such definitions apply to prior as well as future uses of such defined words and phrases.

Brief Description of Drawings

[11] For a more complete understanding of this disclosure and its advantages, reference is now made to the following description taken in conjunction with the accompanying drawings, in which like reference numerals represent like parts:

[12] FIGURE 1 illustrates example scalable video delivery over a heterogeneous network to diverse clients according to this disclosure;

[13] FIGURE 2 illustrates an example MMT device according to this disclosure;

[14] FIGURE 3 illustrates example functionalities provided by Motion Picture Experts Group (MPEG) Media Transport (MMT) according to this disclosure;

[15] FIGURE 4 illustrates an example MMT encapsulation according to this disclosure;

- [16] FIGURE 5 illustrates an example MMT encapsulation with a sample level hint track and a sample level media track of timed media according to this disclosure;
- [17] FIGURE 6 illustrates an example MMT encapsulation of media that does not include timed media according to this disclosure; and
- [18] FIGURE 7 illustrates an example method for receiving Media Fragment Units (MFUs) according to this disclosure.

Best Mode for Carrying out the Invention

- [19] FIGURES 1 through 7, discussed below, and the various embodiments used to describe the principles of this disclosure in this patent document are by way of illustration only and should not be construed in any way to limit the scope of the disclosure. Those skilled in the art understand that the principles of this disclosure may be implemented in any suitable device or system.
- [20] FIGURE 1 illustrates example scalable video delivery over a heterogeneous network to diverse clients according to this disclosure. The embodiment shown in FIGURE 1 is for illustration only. Other embodiments could be used without departing from the scope of this disclosure.
- [21] A heterogeneous network 102 includes a Motion Picture Experts Group (MPEG) Media Transport (MMT) server 104 that serves media content (such as image, audio, and video content) to MMT clients 106-114. The MMT server 104 and the MMT clients 106-114 represent MMT devices and could be implemented as shown in FIGURE 2 or in any other suitable manner. The MMT server 104 can send full-resolution video streams 116 via the heterogeneous network 102 to the MMT clients 106-114. The MMT clients 106-114 receive some or all of the video stream 116 at one or more bit rates 118-126 and one or more resolutions 130-138 based (among other things) on a type of connection to the heterogeneous network 102 and a type of the MMT client. The types and bit rates of connections to the heterogeneous network 102 can include high-speed backbone network connection 128, 1000 megabit per second (Mbps) connection 118, 312 kilobit per second (kbps) connection 120, 1 Mbps connection 122, 4 Mbps connection 124, 2 Mbps connection 126, and so forth. The resolutions 130-138 can include 1080 progressive (1080p) at 60 Hertz (1080p@60Hz) 130, quarter common intermediate format (QCIF) at 10Hz 132, standard definition (SD) at 24Hz 134, 720 progressive (720p) at 60Hz 136, 720p at 30Hz 138, and so on. The types of clients 106-114 can include a desktop computer 106, mobile phone 108, personal digital assistant (PDA) 110, laptop 112, and tablet 114.
- [22] The media content, also referred to as an MMT asset, is sent via MMT containers from the MMT server 104 to the MMT clients 106-114 via one or more MFUs. The MMT asset can be consumed by the MMT clients 106-114, even when not all of the

MFUs are received by the MMT clients 106-114. The MMT container identifies the MFUs that have been received and have not been received so the MMT clients 106-114 can render the MFUs that have been received and not render the MFUs that have not been received.

[23] FIGURE 2 illustrates an example MMT device 202 according to this disclosure. The embodiment of the MMT device 202 shown in FIGURE 2 is for illustration only. Other embodiments of an MMT device could be used without departing from the scope of this disclosure.

[24] The MMT device 202 includes an antenna 205, a radio frequency (RF) transceiver 210, transmit (TX) processing circuitry 215, a microphone 220, and receive (RX) processing circuitry 225. The MMT device 202 also includes a speaker 230, a processing unit 240, an input/output (I/O) interface (IF) 245, a keypad 250, a display 255, and a memory 260. The MMT device 202 could include any number of each of these components.

[25] The processing unit 240 includes processing circuitry configured to execute instructions, such as instructions stored in the memory 260 or internally within the processing unit 240. The memory 260 includes a basic operating system (OS) program 261 and one or more applications 262. The MMT device 202 could represent any suitable device that uses MMT, such as an MMT client 106-114 or an MMT server 104. In particular embodiments, the MMT device 202 represents a mobile telephone, smartphone, personal digital assistant, or tablet computer. The MMT device 202 sends, receives, or displays media content via MMT.

[26] The RF transceiver 210 receives, from the antenna 205, an incoming RF signal transmitted by a base station or other device in a wireless network. The RF transceiver 210 down-converts the incoming RF signal to produce an intermediate frequency (IF) or baseband signal. The IF or baseband signal is sent to the RX processing circuitry 225, which produces a processed baseband signal (such as by filtering, decoding, and/or digitizing the baseband or IF signal). The RX processing circuitry 225 can provide the processed baseband signal to the speaker 230 (for voice data) or to the processing unit 240 for further processing (such as for web browsing or other data).

[27] The TX processing circuitry 215 receives analog or digital voice data from the microphone 220 or other outgoing baseband data (such as web data, e-mail, or interactive video game data) from the processing unit 240. The TX processing circuitry 215 encodes, multiplexes, and/or digitizes the outgoing baseband data to produce a processed baseband or IF signal. The RF transceiver 210 receives the outgoing processed baseband or IF signal from the TX processing circuitry 215 and up-converts the baseband or IF signal to an RF signal that is transmitted via the antenna 205.

[28] In some embodiments, the processing unit 240 includes a central processing unit

(CPU) 242 and a graphics processing unit (GPU) 244 embodied in one or more discrete devices. In some embodiments, the CPU 242 and the GPU 244 are implemented as one or more integrated circuits disposed on one or more printed circuit boards. The memory 260 is coupled to the processing unit 240. In some embodiments, part of the memory 260 represents a random access memory (RAM), and another part of the memory 260 represents a Flash memory acting as a read-only memory (ROM).

[29] In some embodiments, the memory 260 is a computer readable medium that stores program instructions to transfer media content via MMT. When the program instructions are executed by the processing unit 240, the program instructions cause one or more of the processing unit 240, CPU 242, and GPU 244 to execute various functions and programs in accordance with embodiments of this disclosure.

[30] The processing unit 240 executes the basic OS program 261 stored in the memory 260 in order to control the overall operation of the MMT device 202. For example, the processing unit 240 can control the RF transceiver 210, RX processing circuitry 225, and TX processing circuitry 215 in accordance with well-known principles to control the reception of forward channel signals and the transmission of reverse channel signals.

[31] The processing unit 240 is also capable of executing other processes and programs resident in the memory 260, such as operations for identifying MFU locations via a hint track as described in more detail below. The processing unit 240 can also move data into or out of the memory 260 as required by an executing process. In some embodiments, the processing unit 240 is configured to execute a plurality of applications 262. The processing unit 240 can operate the applications 262 based on the OS program 261 or in response to a signal received from a base station. The processing unit 240 is coupled to the I/O interface 245, which provides the MMT device 202 with the ability to connect to other devices, such as laptop computers, handheld computers, and server computers. The I/O interface 245 is the communication path between these accessories and the processing unit 240.

[32] The processing unit 240 is also optionally coupled to the keypad 250 and the display unit 255. An operator of the MMT device 202 uses the keypad 250 to enter data into the MMT device 202. The display 255 may be a liquid crystal display, light emitting diode (LED) display, or other display capable of rendering text and/or at least limited graphics from web sites. Alternate embodiments may use other types of input/output devices and displays.

[33] For efficient and effective delivery of coded media data over heterogeneous Internet Protocol (IP) network environments, MMT provides the following elements:

[34] -a logical model to construct content composed of various components for mash-up applications;

- [35] -a structure of data conveying information about the coded media data for delivery layer processing such as packetization and adaptation;
- [36] -a packetization method and a structure of a packet to deliver media content agnostic to the specific type of media or coding method used over TCP or UDP including hybrid delivery;
- [37] -a format of signaling messages to manage presentation and delivery of media content;
- [38] -a format of signaling messages to manage presentation and delivery of media content; and
- [39] -a format of information to be exchanged across layers to facilitate cross-layer communication.
- [40] FIGURE 3 illustrates example functionalities provided by MMT according to this disclosure. The embodiment shown in FIGURE 3 is for illustration only. Other embodiments could be used without departing from the scope of this disclosure.
- [41] Functionalities provided by MMT are categorized into three functional areas, namely an encapsulation layer 302, a delivery layer 304, and a signaling layer 306. The encapsulation layer 302 defines the logical structure of media content, an MMT package, and a format of the data units to be processed by an MMT-compliant entity. An MMT package includes one or more components having media content and descriptions of relationships among the components to provide information to the underlying delivery layer 304 for adaptive operation. The format of the data units is defined to encapsulate the coded media data of the media content to be stored or carried as a payload of a delivery protocol and to be easily converted between different delivery protocols.
- [42] The delivery layer 304 defines an application layer protocol and a format of the payload. The application layer protocol provides enhanced features, such as multiplexing, for delivery of coded media data compared to conventional application layer protocols for the delivery of media content (also referred to as "multimedia"). The payload format is defined to carry coded media data independent of the specific media type or encoding method.
- [43] The signaling layer 306 defines the format of messages to manage delivery and consumption of MMT packages. Messages for consumption management are used to signal the structure of an MMT package. Messages for delivery management are used to signal the structure of a payload format and a configuration of the protocol.
- [44] The encapsulation layer 302 defines a logical structure and instantiation of an MMT package within an International Organization for Standardization (ISO) Base Media File Format (ISO/BMFF) for distributed storage and adaptive delivery by a network entity. Here, an MMT package defines a logical entity aggregating coded media data of the media content (namely MMT assets) and the information for delivery layer

processing. An MMT asset defines the logical structure carrying the coded media data. MMT assets can collectively reference a number of Media Processing Units (MPUs) with the same MMT asset identification (ID).

- [45] Any type of data that can be individually consumed by an entity directly connected to an MMT client is a separate MMT asset. This includes not only coded media data decodable by a single media codec but also other types of data that have already been multiplexed. MPUs provide information about the media data for adaptive packetization according to the constraints of the underlying delivery layer's packet size, such as the boundaries and sizes of small fragments of the data carried in the MPU. Such small fragments are known as Media Fragment Units (MFUs). This enables the underlying delivery layer entity to dynamically packetize the MPUs adaptively based on the size of the maximum transmission unit of the delivery layer 304. MFUs carry small fragments of coded media data for which such fragments can be independently decoded or discarded, such as a Network Abstraction Layer (NAL) Unit of an Advanced Video Coding (AVC) bitstream.
- [46] An MMT encapsulation format instantiated with ISOBMFF includes several features to correctly capture the logical structure of an MMT package and to achieve the MMT goals and benefits, which include:
- [47] -backwards-compatibility in that MMT encapsulation reuses the ISOBMFF as much as possible and an ISOBMFF parses and accesses/extracts media data from a file storing an MMT package;
- [48] -error robustness to correctly recover an initial structure even in the case of missing MFUs;
- [49] -external addressing to efficiently locate and address MPUs independently by upper layers and the delivery entity without a specific location of the MPUs;
- [50] -support for a unified structure for timed data and non-timed data storage and delivery;
- [51] -ease of access to MPUs and MFUs to easily extract MPU metadata and MFUs for adaptive packetized delivery and to enable easy conversion between storage formats and delivery formats, ideally with the storage format equal to the delivery format;
- [52] -support for encapsulation of legacy data formats, such as MPEG-2 transport stream (TS); and
- [53] -low redundancy.
- [54] Previous MMT designs have several known issues, including:
- [55] -the presence of an MFU "box" for each MFU in the stream, which results in a large amount of MFU boxes and very high overhead (such as an MFU box size of 100 bytes and a NAL size of 1000 bytes, resulting in a 10% overhead);
- [56] -the MFU box references the media data therein, limiting flexibility on how to

- packetize the media content via the delivery layer;
- [57] -the storage format is different from the delivery format, and the delivery format is not clearly defined; and
- [58] -non-timed media data is stored in a manner that is not backwards-compatible.
- [59] In addition, IOSBMFF does not provide various features of MMT, including:
- [60] -independently addressing or consuming movie fragments;
- [61] -correctly parsing and consuming a movie fragment if there are missing parts or packets during packetized delivery; and
- [62] -general usage of hint tracks to support packetization that is adaptive to a size of a maximum transmission unit (MTU) of an underlying delivery network.
- [63] Embodiments of this disclosure encapsulate MMT content within an IOSBMFF by defining a hint track for an MFU structure. Each MFU is described in a hint sample of the hint track of the MMT. The hint sample of the MMT hint track corresponds to an MFU header that is used to deliver an MFU. The MMT hint sample carries information about the location of the corresponding media data. An MMT device uses this information, together with an MFU sequence number, an MPU sequence number, and the asset identifier, to recover the initial MMT asset structure while correctly identifying the MFUs missing from delivery.
- [64] In some embodiments, MMT hint samples are equivalent to MMT Fragmentation Unit headers as illustrated and discussed in relation to FIGURE 4. Some embodiments can also have the MMT hint track samples and the MMT media track samples interleaved on a per sample basis, such as when each MFU is one sample and each chunk includes one sample in both a hint track and a media track, as illustrated and discussed in relation to FIGURES 5 and 6.
- [65] FIGURE 4 illustrates an example MMT encapsulation according to this disclosure. The embodiment shown in FIGURE 4 is for illustration only. Other embodiments could be used without departing from the scope of this disclosure.
- [66] As shown in FIGURE 4, an MMT asset (mast) 402 and an MPU box (mmpu) 404 are delimited by separate boxes. These boxes facilitate external addressability. The storage format shown here is different from the delivery format, and an MMT server uses an MMT hint track 406 to convert the storage format to the delivery format during a packetization operation. The MMT server stores timed media as a track of the IOSBMFF, which contains a single media track 408. The MMT server stores non-timed media as part of the metadata in an IOSBMFF.
- [67] The media track 408 of the IOSBMFF is indexed via a "moov" box 410 in a fully backwards-compatible manner. The MPU box 404 ends at the end of a file or before the start of a subsequent MPU to implicitly indicate the boundaries of an MPU and to allow for more flexibility in editing, recording, and incremental generation of live

content. In some embodiments, the media track 408 of the ISOBMFF is indexed by a "moof" box (such as a moof box 502 in FIGURE 5).

[68] The MMT hint track 406 guides the delivery layer entity, such as an MMT server 104 or MMT clients 106-114, in converting a stored MMT asset 402 into a packetized media stream. The packetized media stream is delivered by an MMT server to an MMT client using the MMT protocol (MMTP) as the delivery protocol.

[69] The MMT hint track 408 provides to the entity generating the delivery packets, such as an MMT server 104, location information of the MFUs of an MPU for adaptive packetization and delivery decisions. The packetization and delivery decisions include statistical multiplexing and dropping of some MFUs to cope with instantaneous network congestion. Media data of an MMT asset within an MMT container are extracted, together with MMT hint samples, to generate MMT packets. One or more MFUs may then be used to build a single MMTP packet. Consequently, the storage formats are not exactly same as the delivery format, which requires an active network entity that is able to extract and encapsulate media data on the fly to build MMTP packets. The media data may be stored in such a way that an ISOBMFF-compliant entity can parse the media data.

[70] The MMT hint track 406 provides information about the MFUs to the network entities for generation of MMT packets. If further fragmentation of an MPU for delivery is not required, the hint track 406 may be omitted completely.

[71] In some embodiments, MMT hint tracks (such as hint track 406) are hint tracks with an entry format in the sample description of "mmth" and can be defined, for example, as:

```
[72] aligned(8) class MMTHintSampleEntry() extends SampleEntry('mmtp' {  
[73]   unsigned int(16) hinttrackversion = 1;  
[74]   unsigned int(16) highestcompatibleversion = 1;  
[75]   unsigned int(1) has_mfus_flag;  
[76]   unsigned int(1) is_timed;  
[77]   unsigned int(6) reserved;  
[78]   MMTAsset asset;  
[79] }
```

[80] The sample description format shown above is for illustration only. Other embodiments could be used without departing from the scope of this disclosure. The semantics of the above sample description format include:

[81] - has_mfus_flag indicates whether the MPUs provide information about MFUs, whether the hint track provides complete MPUs so that each MFU identified by the hint track can have a single sample when the MPUs do not provide information about MFUs, and whether each hint sample refers to an MFU when the MPUs provide in-

formation about MFUs; and

[82] - is_timed: indicates whether the media hinted by this track is timed or is not timed.

[83] Each media sample is assigned to one or more MFUs. Each sample of the MMT hint track can generate one or more MFUs that can use the following syntax:

```
[84] aligned(8) class MMTSample {
[85]   unsigned int(32) sequence_number;
[86]   if (is_timed) {
[87]     signed int(8) trackrefindex;
[88]     unsigned int(32) samplenum;
[89]     unsigned int(16) offset;
[90]     unsigned int(32) length;
[91]   } else {
[92]     unsigned int(16) item_ID;
[93]   }
[94]   unsigned int(1) mfu_lost_flag;
[95]   Unsigned int(7) reserved;
[96] }
```

[97]

[98] The sample format shown above is for illustration only. Other embodiments could be used without departing from the scope of this disclosure. The syntax of the above sample format includes:

[99] - sequence_number is a sequence number assigned to this MFU;

[100] - trackrefindex is an identification of the media track from which the MFU data is extracted;

[101] -samplenum is a sequence number assigned to the sample from which this MFU is extracted;

[102] -item_ID refers to the item that builds this MFU for non-timed media data;

[103] -offset gives the offset of the media data, and the offset base is the beginning of an "mdat" box containing the sample;

[104] -length gives the length of the MFU in bytes; and

[105] - mfu_lost_flag indicates whether the MFU has been lost, such as whether the corresponding media data is missing or not.

[106] FIGURE 5 illustrates an example MMT encapsulation with a sample level hint track and a sample level media track of timed media according to this disclosure. The embodiment shown in FIGURE 5 is for illustration only. Other embodiments could be used without departing from the scope of this disclosure.

[107] In sample-based interleaving for sample delivery, the delivery format corresponds to the MMT payload format and includes the MFU header and the MFU media data. The

MFU header is a copy of the hint sample, which facilitates the creation of the MMTP payload and packets.

[108] In FIGURE 5, a hint track 504 and a media track 506 are interleaved at a sample level inside of a "moof" box 502 of an MMT container 508. This results in one sample per chunk and a straightforward packetization process. The MMT container 508 also includes a "mast" box 510, a "moov" box 512, an "mmpu" box 514, "moof" boxes 502 and 518, and "mdat" boxes 516 and 520. The moof boxes 502 and 518 include MMT hint tracks 504 and 522 and media tracks 506 and 524. The mdat boxes 516 and 520 include one or more MFUs 522. The MFUs 522 include MFU headers 524 and payloads 526. Each MFU header 524 includes an MFU sequence number and optionally includes an `mfu_lost_flag`.

[109] FIGURE 6 illustrates an example MMT encapsulation of media that does not include timed media according to this disclosure. The embodiment shown in FIGURE 6 is for illustration only. Other embodiments could be used without departing from the scope of this disclosure.

[110] In FIGURE 6, a hint track 602 is inside of a "moov" box 604 of an MMT container 606. The MMT container 606 includes a "mast" box 610, the "moov" box 604, an "mmpu" box 614, a "meta" box 632, and an "mdat" box 620. The moov box 604 includes the MMT hint track 602. The meta box 632 includes an "hdlr" box 602, an "iloc" box 630, and an "iinf" box 628. The mdat box 620 includes one or more MFU 622, each of which includes an MFU header 624 and a payload 626.

[111] FIGURE 7 illustrates an example method for receiving MFUs according to this disclosure. The process depicted in FIGURE 7 can be implemented in any suitable manner, such as by the processing circuitry in an MMT server 104, MMT client 106-114, or other device.

[112] In block 702, an MMT device, such as an MMT server 104 or client 106-114, receives an MFU of an MMT container. The MFU is one of one or more MFUs used to transfer media content contained in the MMT container. In block 704, the MMT device splits the MFU into an MFU header and an MFU media data. In block 706, the MMT device uses MPU metadata and an MFU sequence number from the MFU header to locate an MFU hint sample in the MMT container. In some embodiments, the MMT device parses the MFU header, which corresponds to the MFU hint sample.

[113] In block 708, the MMT device determines if the media content of the MMT container includes timed media. If the media content includes timed media, the process continues to block 710. If the media content does not include timed media, the process continues to block 712. In block 710, the MMT device uses MPU metadata and MFU sample information to locate MFU media data from the received MFU in the MMT container. The process then continues to block 714. In block 712, the MMT device

uses MPU metadata and an item identifier to locate media data in the MMT container. The process then continues to block 714.

[114] In block 714, the MMT device marks missing MFUs in MPU metadata in the MMT container. In some embodiments, missing MFUs are marked via the use of an `mfu_lost_flag` value for each MFU within an "mdat" box of the MMT container.

[115] If the MMT asset includes timed media, the MMT device places the MMT hint sample in a correct position within the MMT container during processing. The placement is based on the MFU sequence number, which translates into the sample number in the hint sample. The MMT device also uses the sample and the offset information in the MFU header, together with the sample tables of the media track (such as "stco" "stsz," and "stsc" tables or a "trun" table) to locate the position of the media data in the MMT container. If there are any missing MFUs during the delivery, the corresponding area in the MMT container remains empty. In doing so, the MMT device can safely locate the data for MFUs that are correctly received and skip MFUs that are missing from the delivery.

[116] Although various features have been shown in the figures and described above, various changes may be made to the figures. For example, the size, shape, arrangement, and layout of components shown in the figures are for illustration only. Each component could have any suitable size, shape, and dimensions, and multiple components could have any suitable arrangement and layout. Also, various components in the figures could be combined, further subdivided, or omitted and additional components could be added according to particular needs. Further, each component in a device or system could be implemented using any suitable structure(s) for performing the described function(s). In addition, while FIGURE 7 illustrates various series of steps, various steps in FIGURE 7 could overlap, occur in parallel, occur multiple times, or occur in a different order.

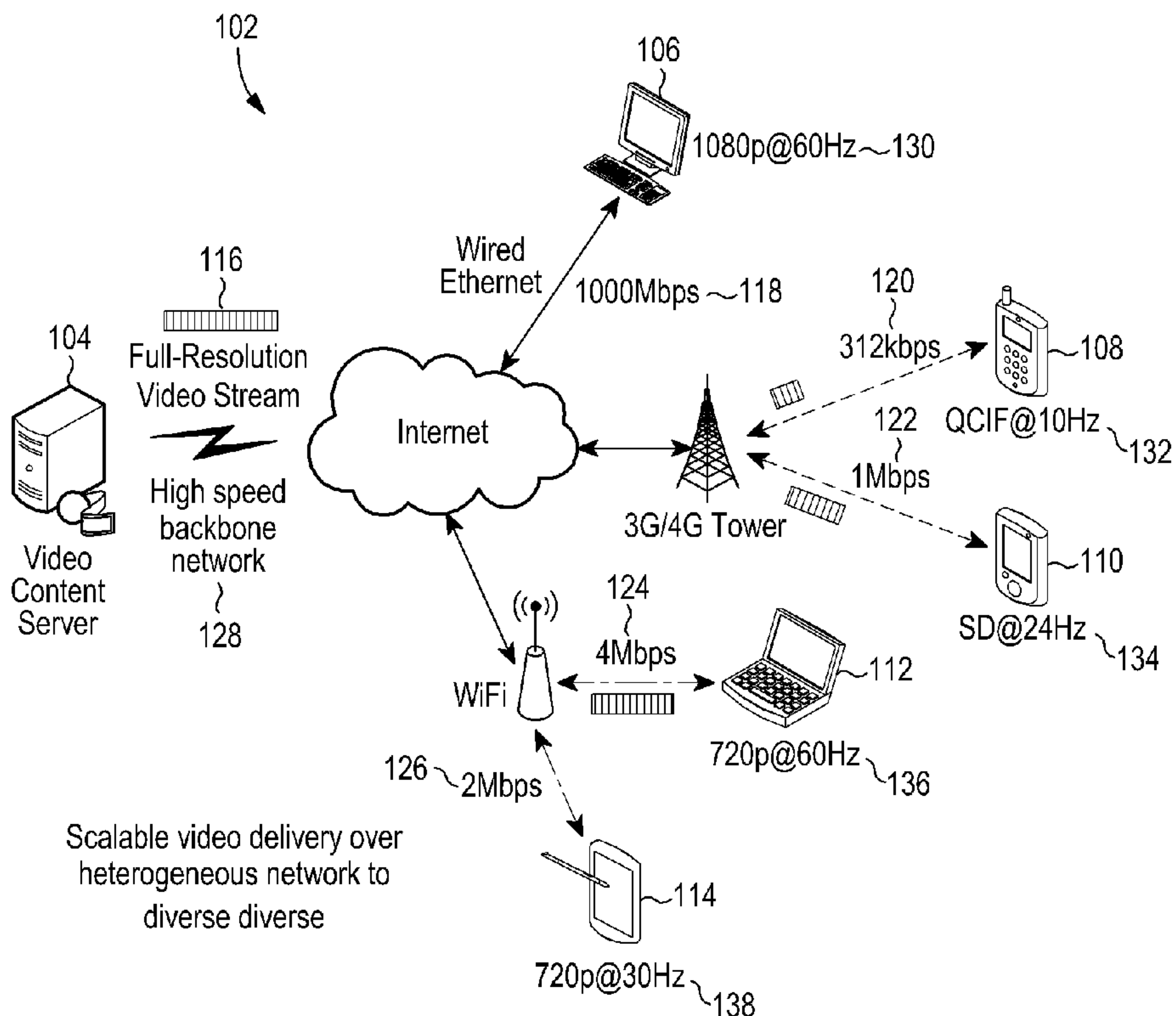
[117] While this disclosure has described certain embodiments and generally associated methods, alterations and permutations of these embodiments and methods are apparent to those skilled in the art. Accordingly, the above description of example embodiments does not define or constrain this disclosure. Other changes, substitutions, and alterations are also possible without departing from the spirit and scope of this disclosure, as defined by the following claims.

Claims

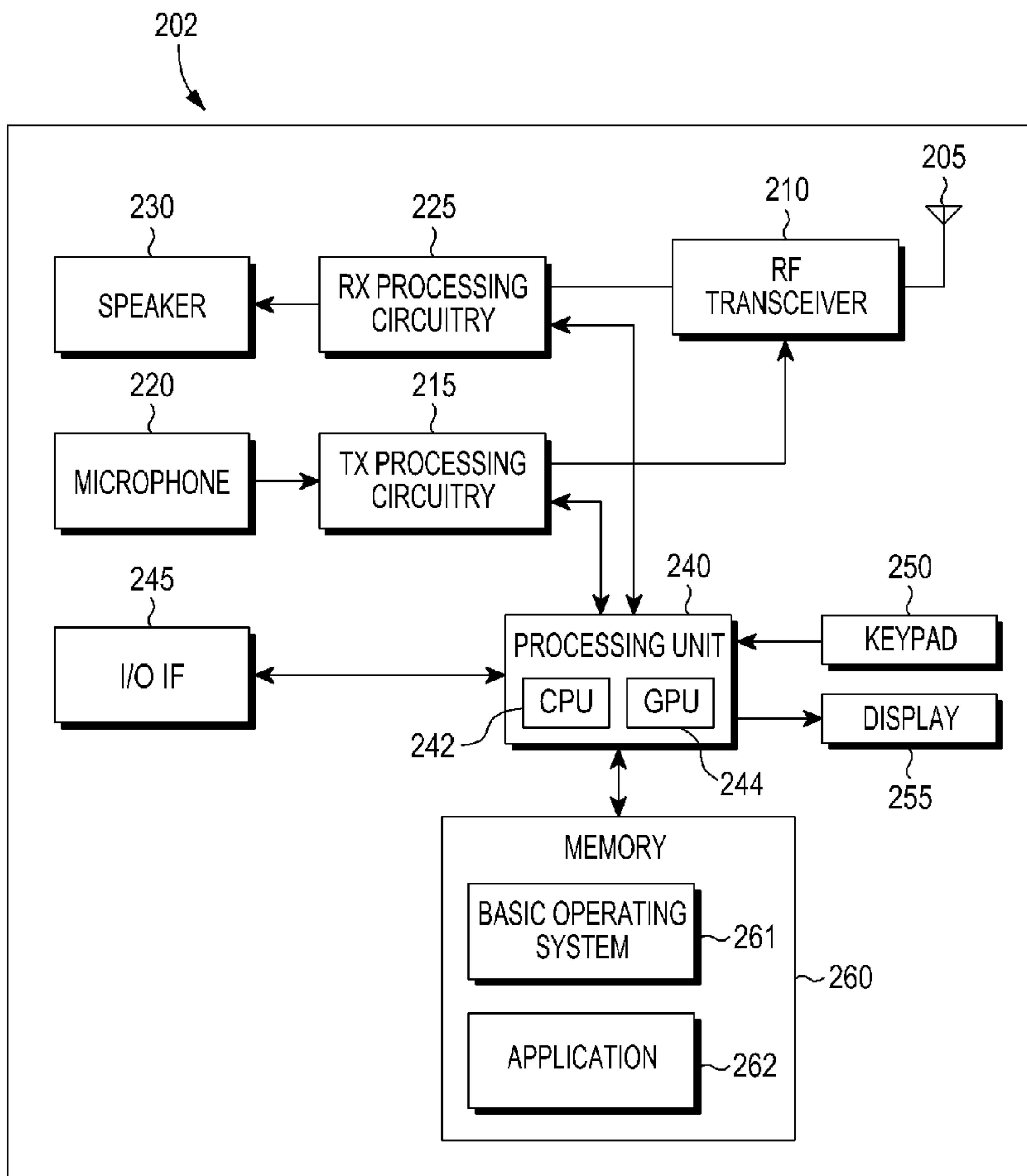
- [Claim 1] An apparatus comprising:
receive path circuitry configured to receive a Motion Picture Experts Group (MPEG) Media Transport (MMT) container; and
a processing device configured to identify locations of one or more media fragment units (MFUs) in the MMT container using a hint track within the MMT container.
- [Claim 2] The apparatus of claim 1, wherein:
the MMT container is an International Organization for Standardization (ISO) Base Media File Format (ISO/BMFF) file; and
the hint track is contained within a moov box of the MMT container.
- [Claim 3] The apparatus of claim 1, wherein:
a fragment of a hint track is contained within a moof box of the MMT container; and
hint samples indexed by the hint track fragment are interleaved on a sample by sample basis.
- [Claim 4] The apparatus of claim 1, wherein a header of each of the one or more MFUs comprises a sequence number of an MFU.
- [Claim 5] The apparatus of claim 4, wherein the header of each of the one or more MFUs comprises an MFU lost flag of an MFU, the MFU lost flag identifying whether the MFU was received by the apparatus.
- [Claim 6] The apparatus of claim 5, wherein, after receiving the one or more MFUs, an MFU lost flag of an MFU that was not received is marked missing.
- [Claim 7] The apparatus of claim 1, wherein:
the MMT container comprises timed media; and
metadata of a media processing unit (MPU) within the MMT container identifies the locations of the one or more MFUs in the MMT container.
- [Claim 8] An apparatus comprising:
transmit path circuitry configured to transmit a Motion Picture Experts Group (MPEG) Media Transport (MMT) container; and
a processing device configured to include a hint track within the MMT container, the hint track identifying locations of one or more media fragment units (MFUs) in the MMT container.
- [Claim 9] The apparatus of claim 8, wherein:
the MMT container is an International Organization for Standardization

- (ISO) Base Media File Format (ISOBMFF) file; and
the hint track is contained within a moov box of the MMT container.
- [Claim 10] The apparatus of claim 8, wherein:
the hint track is contained within a moof box of the MMT container;
and
hint samples indexed by the hint track are interleaved on a sample by
sample basis.
- [Claim 11] The apparatus of claim 8, wherein a header of each of the one or more
MFUs comprises a sequence number of an MFU.
- [Claim 12] The apparatus of claim 11, wherein the header of each of the one or
more MFUs comprises space reserved for an MFU lost flag of an MFU.
- [Claim 13] The apparatus of claim 8, wherein:
the MMT container comprises timed media; and
metadata of a media processing unit (MPU) within the MMT container
identifies the locations of the one or more MFUs in the MMT
container.
- [Claim 14] A method is implemented according to one of claim 1 to 7.
- [Claim 15] A method is implemented according to one of claim 8 to 13.

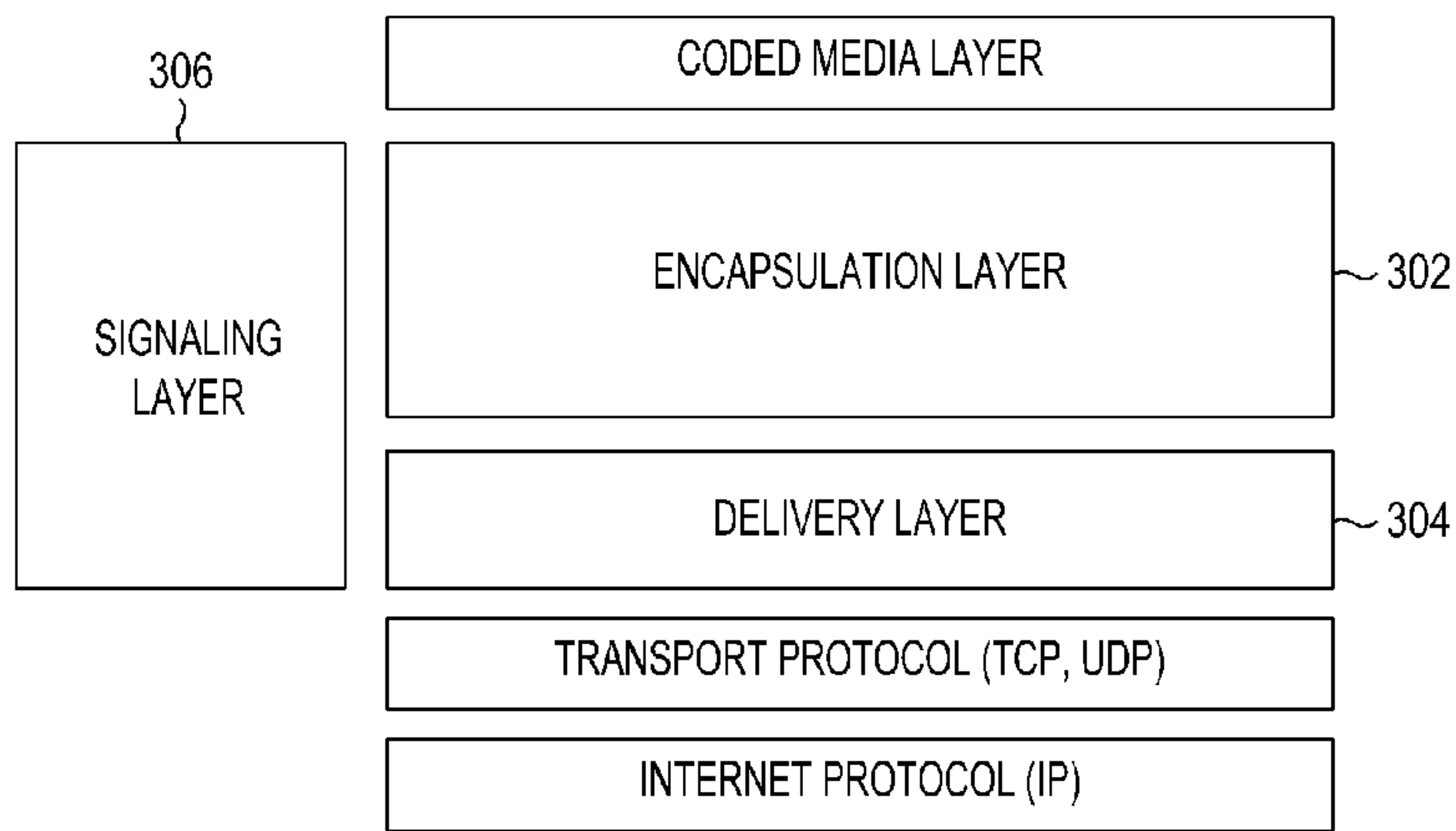
[Fig. 1]



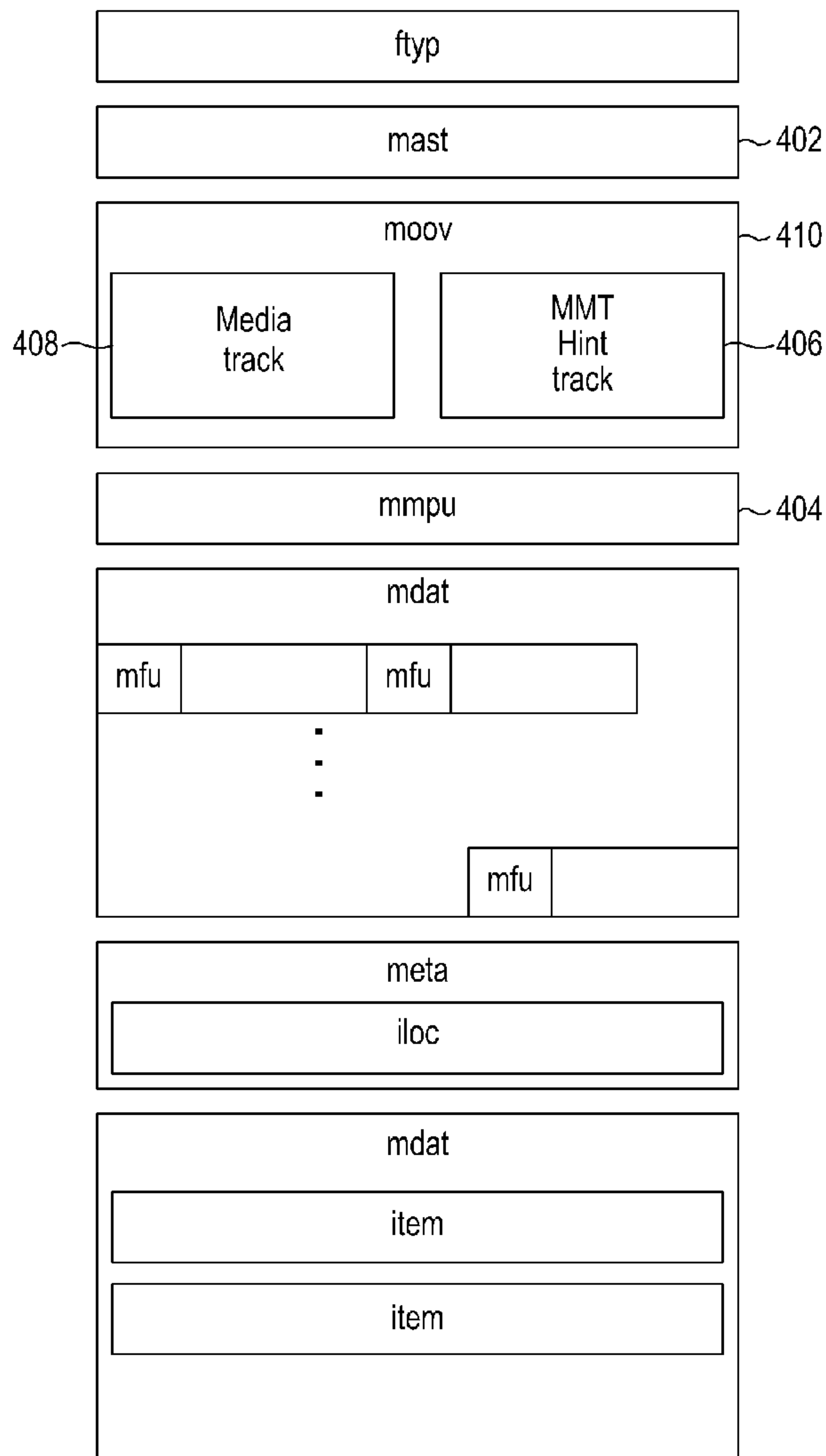
[Fig. 2]



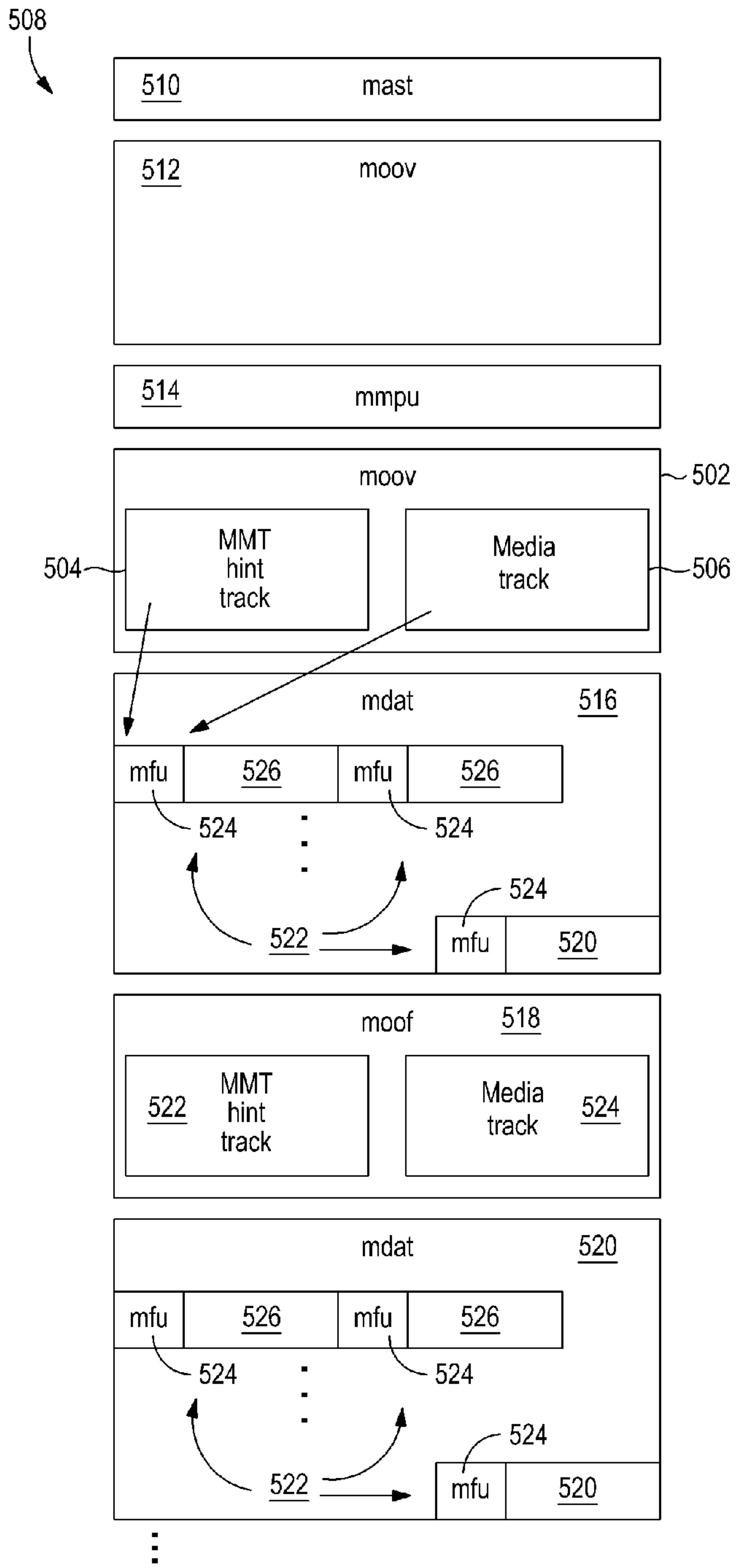
[Fig. 3]



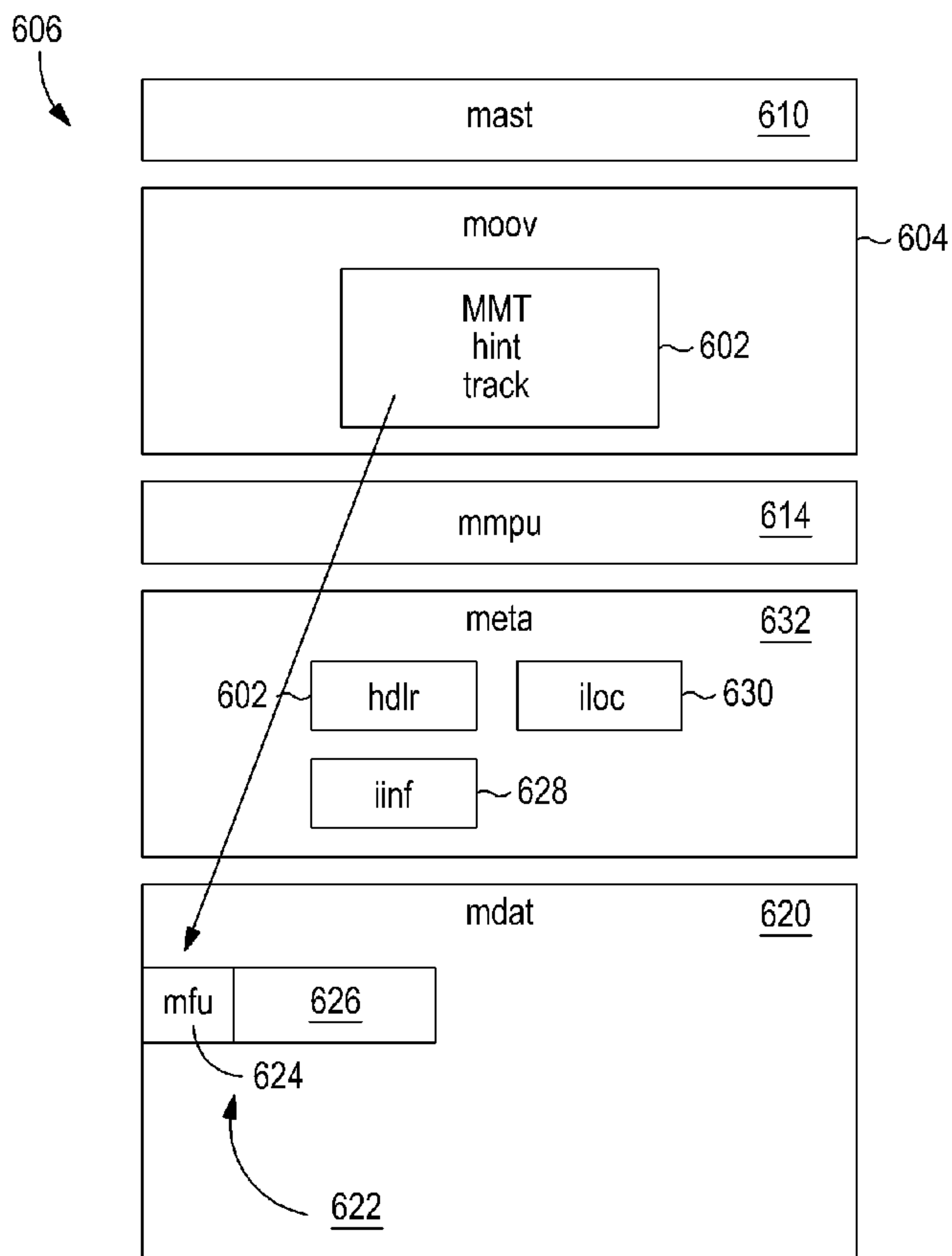
[Fig. 4]



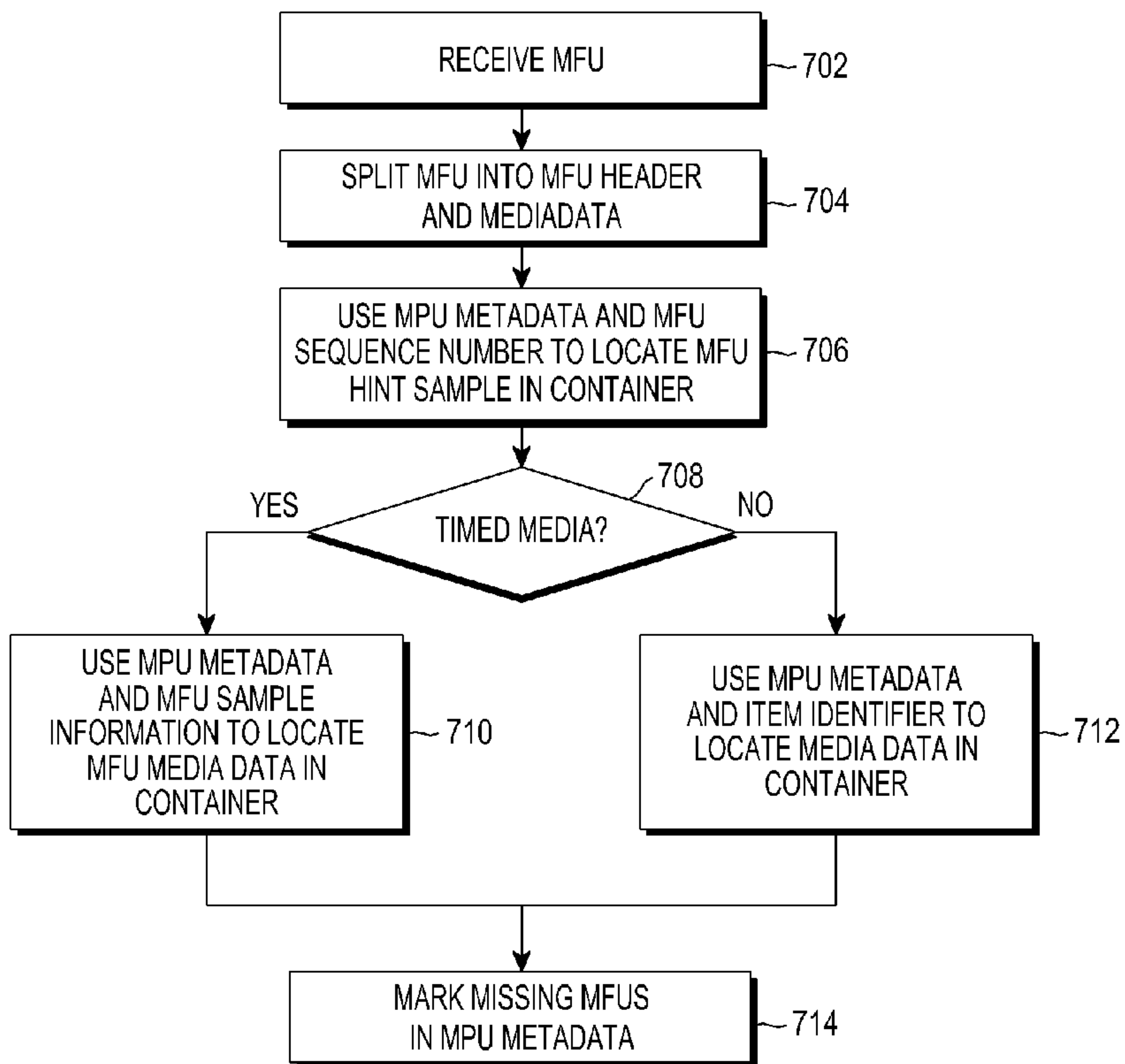
[Fig. 5]



[Fig. 6]



[Fig. 7]



ftyp

mast

402

moov

410

Media track

408

MMT Hint track

406

mmpu

404

mdat

mfu		mfu	
⋮			
			mfu

meta

iloc

mdat

item

item