Office de la Propriété Intellectuelle du Canada
Un organisme d'Industrie Canada

Canadian Intellectual Property Office
An agency of Industry Canada

CA 2868038 A1 2013/10/17

(21) **2 868 038**

(12) **DEMANDE DE BREVET CANADIEN**
CANADIAN PATENT APPLICATION

(13) **A1**

(86) **Date de dépôt PCT/PCT Filing Date:** 2013/04/09
(87) **Date publication PCT/PCT Publication Date:** 2013/10/17
(85) **Entrée phase nationale/National Entry:** 2014/09/19
(86) **N° demande PCT/PCT Application No.:** US 2013/035839
(87) **N° publication PCT/PCT Publication No.:** 2013/155110
(30) **Priorités/Priorities:** 2012/04/09 (US61/621,939); 2012/08/03 (US61/679,627); 2012/09/25 (US13/626,767)

(51) **Cl.Int./Int.Cl.** ***H04N 13/00*** (2006.01), ***H04N 21/23*** (2011.01), ***H04N 21/43*** (2011.01)

(71) **Demandeur/Applicant:**
INTEL CORPORATION, US

(72) **Inventeur/Inventor:**
OYMAN, OZGUR, US

(74) **Agent:** RICHES, MCKENZIE & HERBERT LLP

(54) **Titre : SIGNALISATION D'INFORMATIONS VIDEO TRIDIMENSIONNELLES DANS DES RESEAUX DE COMMUNICATION**
(54) **Title: SIGNALING THREE DIMENSIONAL VIDEO INFORMATION IN COMMUNICATION NETWORKS**

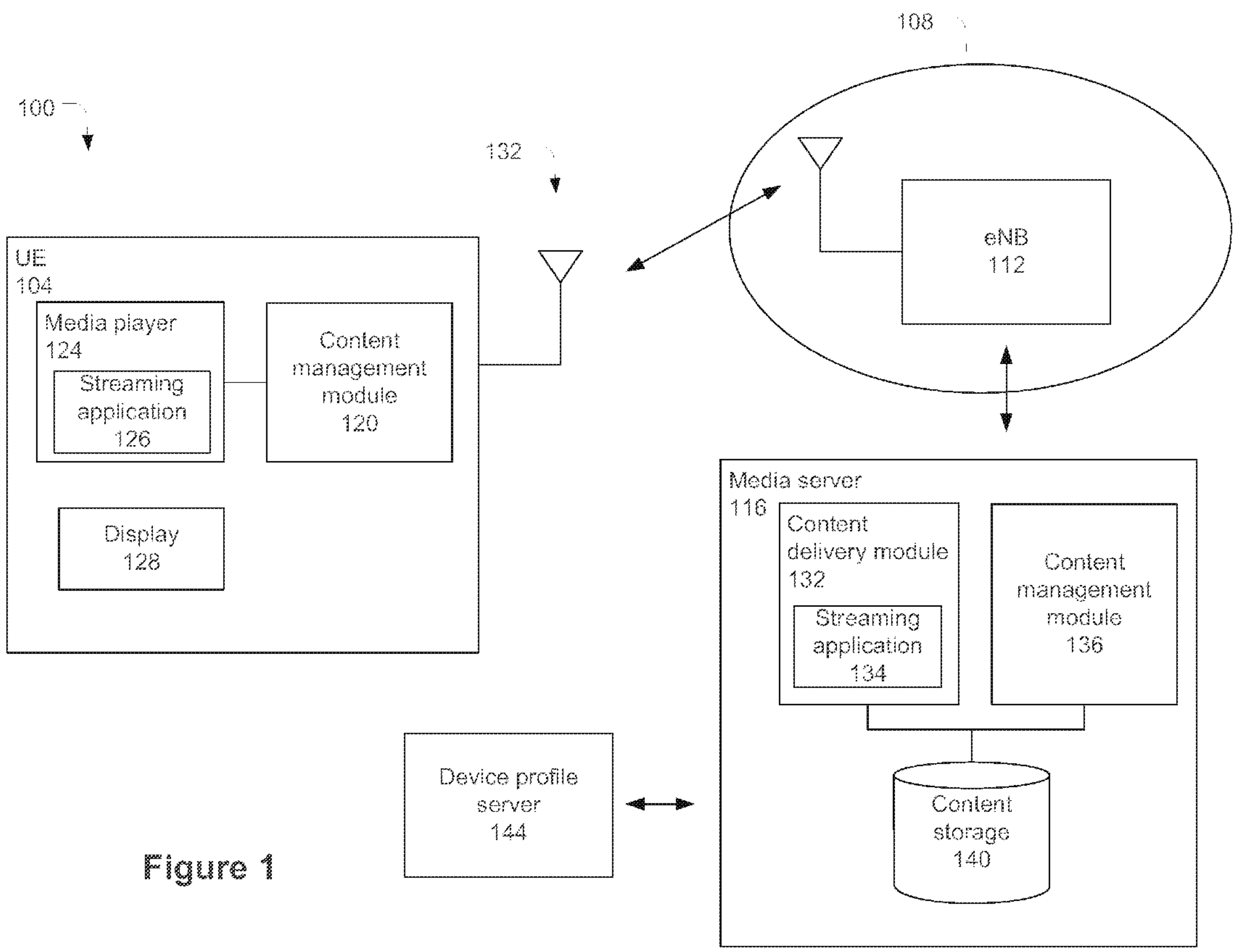

Figure 1

(57) **Abrégé/Abstract:**
Embodiments of the present disclosure describe devices, methods, computer-readable media and systems configurations for signaling stereoscopic three-dimensional video content capabilities of a device in a communications network. Other embodiments may be described and claimed.

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
17 October 2013 (17.10.2013)

WIPO | PCT

(10) International Publication Number
WO 2013/155110 A1

(51) International Patent Classification:
H04N 13/00 (2006.01) *H04N 21/43* (2011.01) *H04N 21/23* (2011.01)

(21) International Application Number: PCT/US2013/035839

(22) International Filing Date: 9 April 2013 (09.04.2013)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/621,939 9 April 2012 (09.04.2012) US
61/679,627 3 August 2012 (03.08.2012) US
13/626,767 25 September 2012 (25.09.2012) US

(71) Applicant *(for all designated States except US)*: **INTEL CORPORATION** [US/US]; 2200 Mission College Blvd, Santa Clara, 95054 (US).

(72) Inventor; and
(71) Applicant *(for US only)*: **OYMAN, Ozgur** [US/US]; 425 Navaro Way, Unit 117, San Jose, California 95134 (US).

(74) Agent: **MAKI, Nathan R.**; Schwabe, Williamson & Wyatt, P.C., 1211 SW 5th Avenue, Suite 1900, Portland, Oregon 97204 (US).

(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, IS, JP, KE, KG, KM, 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, 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, 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, ML, MR, NE, SN, TD, TG).

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

(54) **Title**: SIGNALING THREE DIMENSIONAL VIDEO INFORMATION IN COMMUNICATION NETWORKS

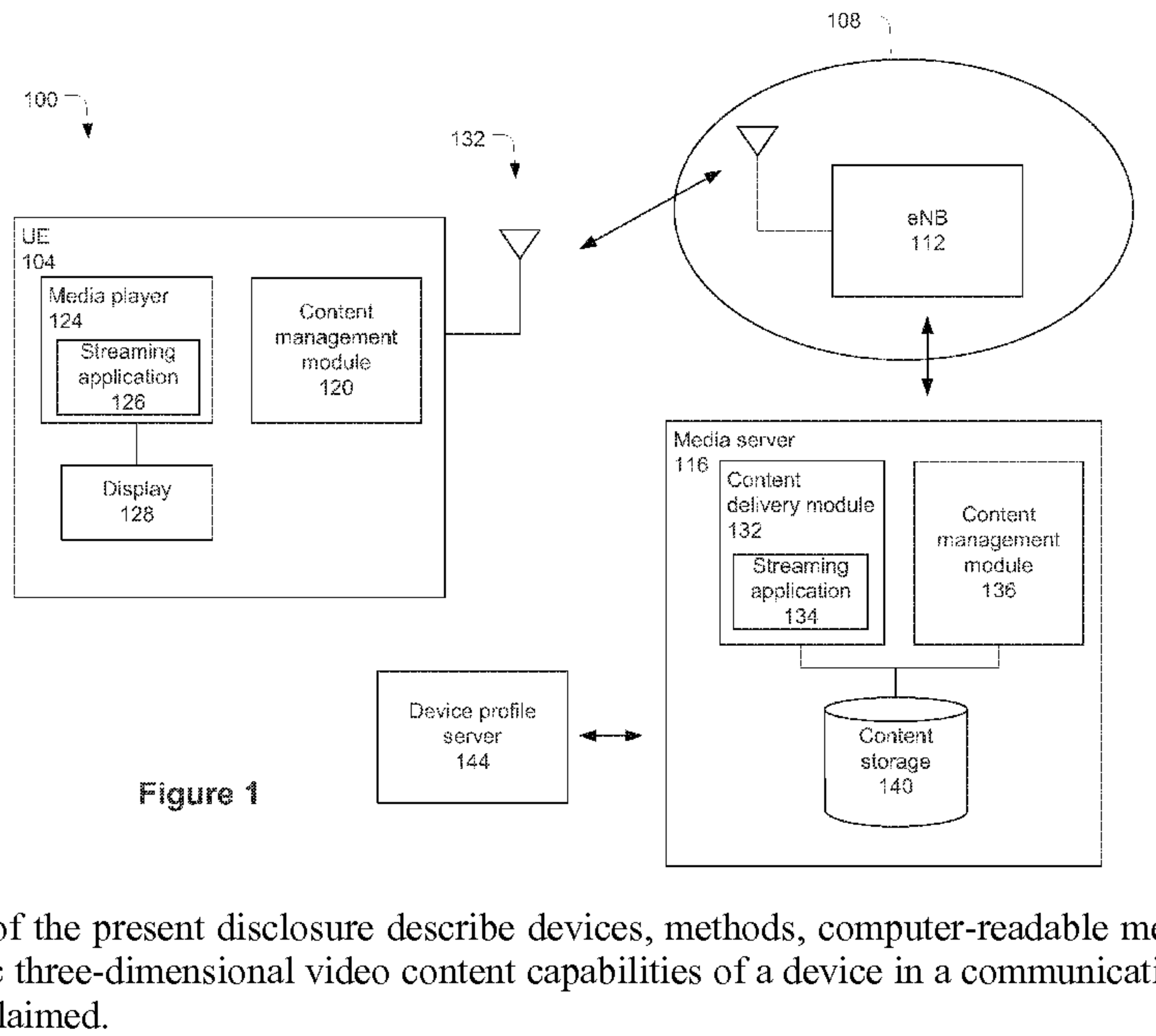

Figure 1

(57) **Abstract**: Embodiments of the present disclosure describe devices, methods, computer-readable media and systems configurations for signaling stereoscopic three-dimensional video content capabilities of a device in a communications network. Other embodiments may be described and claimed.

WO 2013/155110 A1

SIGNALING THREE DIMENSIONAL VIDEO INFORMATION IN COMMUNICATION NETWORKS

Field

Embodiments of the present invention relate generally to the field of communications, and more particularly, to signaling three-dimensional video information in communication networks.

Background

Three-dimensional (3-D) video offers a high-quality and immersive multimedia experience, which has only recently become feasible on consumer electronics and mobile platforms through advances in display technology, signal processing, transmission technology, and circuit design. It is currently being introduced to the home through various channels, including by Blu-ray Disc™, cable and satellite transmission, etc., as well as to mobile networks through 3-D enabled smartphones, etc. Concepts related to delivery of such content through wireless networks are being developed.

Brief Description of the Drawings

Embodiments will be readily understood by the following detailed description in conjunction with the accompanying drawings. To facilitate this description, like reference numerals designate like structural elements. Embodiments are illustrated by way of example and not by way of limitation in the figures of the accompanying drawings.

Figure 1 schematically illustrates a wireless communication network in accordance with various embodiments.

Figures 2a-b illustrate adaptation of streamed content and/or associated session description and metadata files in accordance with various embodiments.

Figure 3 illustrates a setup of a streaming session in accordance with an embodiment.

Figure 4 illustrates frame compatible packing formats in accordance with various embodiments.

Figure 5 illustrates a method of signaling 3-D video device capabilities in accordance with various embodiments.

Figure 6 illustrates a method of signaling 3-D video content in accordance with various embodiments.

Figure 7 schematically depicts an example system in accordance with various embodiments.

Detailed Description

Illustrative embodiments of the present disclosure include, but are not limited to, methods, systems, computer-readable media, and apparatuses for signaling stereoscopic three-dimensional video content capabilities of a client device in a communication network. Some embodiments of this invention in this context could be on methods, systems, computer-readable media, and apparatuses for signaling stereoscopic three-dimensional video content capabilities of a mobile device in a wireless communications network.

Various aspects of the illustrative embodiments will be described using terms commonly employed by those skilled in the art to convey the substance of their work to others skilled in the art. However, it will be apparent to those skilled in the art that alternate embodiments may be practiced with only some of the described aspects. For purposes of explanation, specific numbers, materials, and configurations are set forth in order to provide a thorough understanding of the illustrative embodiments. However, it will be apparent to one skilled in the art that alternate embodiments may be practiced without the specific details. In other instances, well-known features are omitted or simplified in order not to obscure the illustrative embodiments.

Further, various operations will be described as multiple discrete operations, in turn, in a manner that is most helpful in understanding the illustrative embodiments; however, the order of description should not be construed as to imply that these operations are necessarily order dependent. In particular, these operations need not be performed in the order of presentation.

The phrase "in some embodiments" is used repeatedly. The phrase generally does not refer to the same embodiments; however, it may. The terms "comprising," "having," and "including" are synonymous, unless the context dictates otherwise. The phrase "A and/or B" means (A), (B), or (A and B). The phrases "A/B" and "A or B" mean (A), (B), or (A and B), similar to the phrase "A and/or B". The phrase "at least one of A, B and C" means (A), (B), (C), (A and B), (A and C), (B and C) or (A, B and C). The phrase "(A) B" means (B) or (A and B), that is, A is optional.

Although specific embodiments have been illustrated and described herein, it will be appreciated by those of ordinary skill in the art that a wide variety of alternate and/or equivalent implementations may be substituted for the specific embodiments shown and described, without departing from the scope of the embodiments of the present disclosure. This application is intended to cover any adaptations or variations of the embodiments discussed herein. Therefore, it is manifestly intended that the embodiments of the present disclosure be limited only by the claims and the equivalents thereof.

As used herein, the term "module" may refer to, be part of, or include an Application Specific Integrated Circuit (ASIC), an electronic circuit, a processor (shared, dedicated, or group) and/or memory (shared, dedicated, or group) that execute one or more software or firmware programs, a combinational logic circuit, and/or other suitable components that provide the described functionality

Significant improvements in video compression capability have been demonstrated with the introduction of the H.264/MPEG-4 advanced video coding (AVC) standard. Since developing the standard, the joint video team of the ITU-T Video Coding Experts Group (VCEG) and the International Organization for Standardization (ISO)/International Electrotechnical Commission (IEC) Moving Picture Experts Group (MPEG) has also standardized an extension of AVC that is referred to as multiview video coding (MVC). MVC provides a compact representation for multiple views of the video scene, such as multiple synchronized video cameras.

In stereoscopic 3-D video applications, two views are displayed. One for the left eye and one for the right eye. There are various ways of formatting the views of stereoscopic 3-D video content. In one embodiment, the encoding of stereo-paired 3-D video may be a special case of MVC, where left and right eye views are produced via MVC. Other encoding formats of producing 3-D video content are also possible. Various devices may have different capabilities with respect to decoding and rendering these different formats. Embodiments described herein provide for various parameters of a device capability exchange that may facilitate delivery and viewing of the 3-D video content in a communication network such as a wireless network, e.g., an evolved universal terrestrial radio access network (EUTRAN).

Figure 1 schematically illustrates a network environment 100 in accordance with various embodiments. The network environment 100 includes a user equipment (UE) 104, which may also be referred to as a client terminal or mobile device, wirelessly coupled with a radio access network (RAN) 108. The RAN 108 may include an enhanced node base station (eNB) 112 configured to communicate with the UE 104 via an over-the-air (OTA) interface. The RAN 108 may be part of a third generation partnership project (3GPP) long-term evolution (LTE) advanced network and may be referred to as an EUTRAN. In other embodiments, other radio access network technologies may be utilized.

The UE 104 may communicate with a remote media server 116 through the RAN 108. While the eNB 112 is shown communicating directly with the media server, it will be understood that the communications may flow through a number of intermediate networking components, e.g., switches, routers, gateways, etc., in various embodiments. For example, in some embodiments, the RAN 108 may be coupled with a core services network (CSN) that communicatively couples the RAN 108 with a larger network, e.g., a wide area network, of which the media server 116 may be considered a part.

While Figure 1 describes the network environment as a wireless communication network, other embodiments may be used in other types of networks, e.g., wire-line networks. It may be understood that other network environments in which embodiments of the present invention may be employed may include additional, fewer, or different components than those explicitly shown in the example depicted in Figure 1. For example, embodiments of the present invention employed in a wire-line network, may have the media server 116 and the UE 104 communicating with one another without the RAN 108.

The UE 104 and media server 116 may have a number of components that are configured to facilitate access, storage, transmission, and display of 3-D video content. For example, the UE 104 may include a content management module 120, a media player 124 having a streaming application 126, and a display 128. The streaming application 126 may have sufficient functionality to receive 3-D video content and associated information; decode, unpack, and otherwise re-assemble the 3-D video; and render the 3-D video on the display 128. In various embodiments, the streaming application 126 may be referred to in the context of the streaming technology employed. For example, in embodiments in which the content is streamed by a packet-switched streaming service (PSS), the streaming application 126 may be referred to as a PSS application. The content management module 120 may negotiate or otherwise communicate streaming parameters including, e.g., device capability parameters, to enable receipt of the data in manner that facilitates operation of the media player 124.

The media server 116 may include content delivery module 132 having a streaming application 134, a content management module 136, and a content storage 140. The content delivery module 132 may encode, pack, or otherwise assemble 3-D video content, stored in the content storage 140, for transmission to one or more UEs, e.g., UE 104. The content management module 136 may negotiate or otherwise communicate streaming parameters including, e.g., device capability parameters, and control the content delivery module 132 in a manner to facilitate delivery of the 3-D content.

In some embodiments, one or more of the components that are shown as being part of the media server 116 may be disposed separately from the media server 116 and communicatively coupled with the media server over a communication link. For example, in some embodiments, content storage 140 may be disposed remotely from the content delivery module 132 and the content management module 136.

In some embodiments, the content delivery module 132 may deliver, through eNB 112 in one example, the 3-D video content to the UE 104 in accordance with a 3GPP streaming standard. For example, the 3-D video content may be transmitted in accordance with a PSS standard, e.g., 3GPP TS 26.234 V11.0.0 (March 16, 2012), a dynamic adaptive streaming over HTTP (DASH) standard, e.g., 3GPP TS 26.247 V.11.0.0 (March 16, 2012), a multimedia broadcast and multicast service (MBMS) standard, e.g., TS 26.346 V11.1.0 (June 29, 2012), and/or an IMS-based PSS and MBMS services (IMS_PSS_MBMS) standard, e.g., TS 26.237 V.11.0.0 (June 29, 2012). The streaming application 126 may be configured to receive the 3-D video content over any of a number of transport protocols, e.g., real-time transport protocol (RTP), hypertext transport protocol (HTTP), etc.

Capability exchange enables media streaming servers, such as media server 116, to provide a wide range of devices with video content suitable for the particular device in question. To facilitate server-side content negotiation for streaming, the media server 116 may determine the specific capabilities of the UE 104.

The content management module 120 and the content management module 136 may negotiate or otherwise communicate parameters of a 3-D video content streaming session. This negotiation may take place through session-level signaling via the RAN 108. In some embodiments, the session-level signaling may include transmissions related to device capability information that includes stereoscopic 3-D video decoding and rendering capabilities of the media player 124. In various embodiments, the device capability information may further include pre-decoder buffer size, initial buffering., decoder capability, display properties (screen size, resolution, bit depth, etc.), streaming method (real-time streaming protocol (RTSP), HTTP, etc.) adaptation support, quality of experience (QoE) support, extended real-time transport protocol (RTCP) reporting support, fast content switching support, supported RTP profiles, session description protocol (SDP) attributes, etc.

During the setup of the streaming session, the content management module 136 may use the device capability information to control the content delivery module 132 in a manner to provide the UE 104 with the proper type of multimedia content. For example, the media server 116 may determine which variants of multiple available variants of a video stream are desired based on the actual capabilities of the UE 104 to determine the best-suited streams for that terminal. This may allow for improved delivery of 3-D video content and associated session description and metadata files, for example SDP file or a media presentation description (MPD) file, to the UE 104.

The content delivery module 132 may access the content in the content storage 140 and adapt the content and/or associated session description and metadata files, e.g., SDP/MPD files, according to the negotiated session parameters prior to delivery of the content/associated files. The content, when delivered to the UE 104, may be decoded by the media player 124 and rendered on the display 128.

Adaptation of content and/or associated session description and metadata files is shown in accordance with some specific examples with reference to Figures 2a-b, while setup of streaming session is shown in accordance with a specific example with reference to Figure 3.

Figure 2a illustrates a DASH-based streaming embodiment with adaptation of 3-D video formats in accordance with some embodiments. In particular, Figure 2a illustrates an HTTP server 204 in communication with a DASH client 208 and implementing a pull-based streaming embodiment, in which the streaming control is maintained by the client rather than the server, where the client downloads content from the server through a series of HTTP-based request-response transactions after the inspection of the MPD. In DASH-based streaming, the MPD metadata file provides information on the structure and different versions of the media content representations stored in the HTTP server 204 (including different bitrates, frame rates, resolutions, codec types, etc.). Based on this MPD metadata information that describes the relation of the segments and how they form a media presentation, DASH client 208 may request the media segments using HTTP GET or partial GET methods. The HTTP server 204 and DASH client 208 may be similar to and substantially interchangeable with media server 116 and UE 104, respectively.

In DASH, the set of 3-D video formats and corresponding content information may be signaled to the DASH client 208 in the MPD. Depending on the capability profile of the DASH client 208 and its supported 3-D formats, the HTTP server 204 may offer different formatted content, e.g., the HTTP server 204 may exclude the 3-D formats that are not supported by the DASH client 208 in the MPD and only include those that are supported by the DASH client 208. In this context, the HTTP server 204 may provide the content optimized for different 3-D video formats to the DASH client 208. In doing this, the HTTP server 204 may use the device capability exchange signaling from the DASH client 208 describing the various supported 3-D video formats. The DASH client 208 may then request the corresponding versions of the 3-D video content supported by the DASH client 208. Moreover, when retrieving an MPD with HTTP, the DASH client 208 may include 3-D video codec and format information in a GET request, including any temporary adjustments to the 3-D video formats based on profile difference (ProfDiff). In an example, the difference may be configured to temporarily modify one or more MPD parameters for a content presentation session. For example, the difference may be configured to modify the MPD until the content presentation session ends or a subsequent difference (corresponding to the first communicated difference) is communicated to the HTTP server 204. This way the HTTP server 204 may deliver an optimized MPD to the DASH client 208.

Figure 2b illustrates an RTSP-based streaming embodiment with adaptation of 3-D video formats in accordance with some embodiments. In particular, Figure 2b illustrates a server 212 and a client 216 implementing a push-based streaming method, in which the streaming and session control are maintained by the server 212 rather than the client 216. The server 212 and client 216 may be similar to and substantially interchangeable with media server 116 and UE 104, respectively.

Examples of push-based streaming include PSS and IMS_PSS_MBMS services based on the RTSP and session initiation protocol (SIP), respectively. In this context, the server 212 receives the set of supported 3-D video codecs and formats from the client 216 and adapts the content based on this information, e.g., the server 212 selects the most suited content version among stored content versions or dynamically transcodes the content based on the supported 3-D video formats and streams the content to the client 216. The session-related metadata carried in the SDP may carry the 3-D video format information for the streamed content.

Figure 3 illustrates a service discovery with subscribe/notify for IMS_PSS_MBMS service in accordance with some embodiments. In particular, Figure 3 illustrates interactions between a UE 304, an IP Multimedia (IM) Core Network (CN) subsystem 308, and a service discovery function (SDF) 312. The UE 304 may be similar to and substantially interchangeable with UE 104. The IM CN subsystem 308 and the SDF 312 may be part of a core network domain that interfaces with the access network domain, e.g., the RAN 108.

In the IMS_PSS_MBMS service, the UE 304 can send device capability information, e.g., supported 3-D video codecs and formats, in a SIP SUBSCRIBE message to the IM CN Subsystem 308 during service discovery. The IM CN subsystem 308 may then forward the message to the SDF 312. The SDF 312 determines the proper service discovery information, e.g. according to the capabilities of the UE 304 as described in the user's profile (Personalized Service Discovery). The SDF 312 may then send a SIP 200 OK message to the

IM CN subsystem 308, which is relayed to the UE 304 to confirm the session initialization based on the sent device capability information that also includes the supported 3-D video codecs and formats. Afterward, the SDF 312 may send a SIP NOTIFY message, with service discovery information, to the IM CN subsystem 308, which relays the SIP NOTIFY message back to the UE 304. The UE 304 may then respond by sending a SIP 200 OK message to the IM CN subsystem 308, which is then relayed to the SDF 312.

Such a framework enables optimized service discovery utilizing the supported 3-D video formats in IMS-based PSS and MBMS user services. Later during the IMS session, the UE 304 may also use SIP signaling to indicate updates including any temporary adjustments to the set of supported 3-D video codecs and formats based on ProfDiff (e.g., if the current device orientation is different from the default device orientation). This may be done by refreshing the subscription through further SIP SUBSCRIBE messages including information on the updates to the 3-D video format information.

Referring again to Figure 1, in some embodiments, the media server 116 may be coupled with a device profile server 144 that has profile information of the UE 104. The profile information may include some or all of the device capability information. In such embodiments, the media server 116 may receive identification information from the UE 104 and then retrieve the profile information from the device profile server 144. This may be done as part of the session-level signaling.

In some embodiments, the UE 104 may supplement the profile information retrieved from the device profile server 144 with extra attributes or overrides for attributes already defined in its device capability profile, based on ProfDiff signaling. In one example, such a temporary adjustment may be triggered by user preferences, for example if the user for a particular session only would like to receive two-dimensional (2-D) video even though the terminal is capable of rendering 3-D video.

The streaming application 134 may encode the 3-D video content for transmission in the network environment 100 in accordance with a number of different stream types, with each stream type having associated frame types. Frame types could include frame packing, simulcast, or 2-D plus auxiliary frame types.

Frame packing may include frame-compatible packing formats and full-resolution per view (FRPV) packing format. In frame-compatible packet formats, the streaming application 134 may spatially pack constituent frames of a stereo pair into a single frame and encode the single frame. Output frames produced by the streaming application 126 contain constituent frames of a stereo pair. The spatial resolution of the original frames of each view and the packaged single frame may be the same. In this case, the streaming application 134 may down-sample the two constituent frames before the packing operation. The frame-compatible packing formats may use a vertical interleaving, horizontal interleaving, side-by-side, top-bottom, or checkerboard format as illustrated in Figures4a-e, respectively, and the down sampling may be performed accordingly.

In some embodiments, the streaming application 134 may indicate the frame-packing format that was used by including one or more frame packing arrangement supplemental enhancement information (SEI) messages as specified in the H.264/AVC standard into the bitstream. The streaming application 126 may decode the frame, unpack the two constituent frames from the output frames of the decoder, up sample the frames to revert the encoder side down sampling process, and render the constituent frames on the display 128.

A FRPV packing format may include temporal interleaving. In temporal interleaving, the 3-D video may be encoded at double the frame rate of the original video with each parent and subsequent pictures constituting a stereo pair (left and right view). The rendering of the time interleaved stereoscopic video may typically be performed at a high frame rate, where active (shutter) glasses are used to blend the incorrect view in each eye. This may rely on accurate synchronization between the glasses and the screen.

In embodiments using simulcast frame types, the left and the right views may be transmitted in separate, simulcast streams. The separately transmitted streams may be combined by the streaming application 126 and jointly decoded.

In embodiments using 2-D plus auxiliary frame types, 2-D video content may be sent by the streaming application 134 in conjunction with auxiliary information that may be used by the streaming application 126 to render 3-D video on the display 128. This auxiliary information may be, e.g., a depth/parallax map that is a 2-D map with each pixel defining a depth/parallax of one or more pixels in an associated 2-D video frame.

In some embodiments, other frame types may be used. For example, in some embodiments the streaming application 134 may be capable of encoding stereoscopic views into a base view stream and a non-base view stream, which may be transmitted in the same or different streams. In some embodiments, this may be referred to as MVC-based for stereoscopic video. The non-base view stream may include inter-view prediction frames that provide spatial/temporal predictive information. The base view stream may be sufficient for a single-view, e.g., 2-D, decoder to render the base view as 2-D video, while the non-base view stream may provide 3-D decoders, e.g., streaming application 126, with sufficient information to render 3-D video. If the media server 116 is aware of UEs' capabilities, it can omit sending the non-base view stream to a device that does not support 3-D video or does not have sufficient bitrate to support 3-D video.

In various embodiments, the device capability information, transmitted from content management module 120 and/or device profile server 144 to content management module 136, may include a 3-D format attribute that includes a list of one or more formats relevant for streaming of stereoscopic 3-D video over relevant transmission protocol, e.g., RTP or HTTP, supported by the streaming application 126. In some embodiments, the 3-D format attribute may be a streaming frame packing format for RTP or HTTP having an integer value "1" for vertical interleaving, "2" for horizontal interleaving, "3" for side-by-side, "4" for top-bottom, "0" for checkerboard, or "5" for temporal interleaving. In some embodiments, the same 3-D format attributes may be used to indicate frame packing formats supported in a specific file or container format. In some embodiments, the 3-D format attribute may include a more generalized value, e.g., "FP" for frame packing.

In some embodiments, the 3-D format attribute may be another streaming format having a value "SC" for simulcast or "2DA" for 2-D video plus auxiliary information.

In embodiments in which the UE 104 supports more than one format type, it may further indicate one or more preferred format types. This could be done by listing the format types in an order of preference, associating a preference indicator with select format types, etc.

In some embodiments, in addition to providing a frame type attribute, the content management module 120 and/or the device profile server 144 may provide one or more component type attributes. The component type attributes may provide additional details about specific types of video components, which are constituent elements of the stereoscopic 3-D video, supported and/or preferred by the streaming application 126.

The component type attributes may have a value "C" for indicating a center-view stream, "CD" for indicating a center-view stream and a depth map, "CP" for indicating a center-view stream and a parallax map, "D" for indicating a depth map, "P" for indicating a parallax map, "L" for indicating a left-view stream, "LD" for indicating a left-view stream and a depth map, "LIL" for indicating video frames that include alternating scan lines from the left and right views, "LP" for indicating a left-view stream and a parallax map, "R" for indicating a right-view stream, "Seq" to indicate frame sequential (e.g., video stream that includes alternating frames from the left and right streams – additional signaling, e.g., AVC SEI messages, may be needed to signal which frames contain left and right views), "SbS" for indicating side-by-side, and "TaB" for indicating top and bottom.

Each format type attribute may be associated with a respective set of component type attributes. For example, if the format type is SC, the associated component type may be L or R to indicate left and right views, respectively.

The device capability exchange signaling capability in the PSS specification 3GPP TS 24.234 enables servers to provide a wide range of devices with content suitable for the particular device in question. In order to improve delivery of stereoscopic 3-D video content to the client terminal, the present disclosure describes a new set of attributes that may be included in the PSS vocabulary for device capability exchange signaling. These proposed attributes may describe the 3-D video decoding and rendering capabilities of the client terminal, including which 3-D video frame packing formats the client supports. This may for example allow the server and network to provide an optimized RTSP SDP or DASH MPD to the client terminal, as well as to perform the appropriate transcoding and 3-D format conversions in order to match the transmitted 3-D video content to the capabilities of the client device.

The device capability exchange signaling of supported 3-D video codecs and formats may be enabled in 3GPP TS 26.234 with the inclusion of three new attributes in the PSS vocabulary: (1) for Streaming component, two attributes indicating the list of supported frame packing formats relevant for streaming of stereoscopic 3-D video over RTP and HTTP, respectively, and (2) for ThreeGPFileFormat component, one attribute indicating the list of supported frame packing formats relevant for stereoscopic 3-D video that can be included in a 3GPP file format (3GP) file, which is a multimedia container format commonly used for 3GPP-based multimedia services. The details of the attribute definitions are presented below in accordance with some embodiments.

Attribute name: StreamingFramePackingFormatsRTP

Attribute definition: List of supported frame packing formats relevant for streaming of stereoscopic 3-D video over RTP supported by the PSS application. The frame packing formats within scope for stereoscopic 3-D video include:

Frame Compatible Packing Formats: 1 = Vertical interleaving, 2 = Horizontal interleaving. 3 = Side-by-Side, 4 = Top-Bottom, 0 = Checkerboard

Full-Resolution per View Packing Formats: 5 = Temporal Interleaving

Component: Streaming

Type: Literal (Bag)

Legal values: List of integer values corresponding to the supported frame packing formats.

Resolution rule: Append

EXAMPLE: <StreamingFramePackingFormatsRTP>
<rdf:Bag>
<rdf:li>3</rdf:li>
<rdf:li>4</rdf:li>
</rdf:Bag>
</StreamingFramePackingFormatsRTP>

Attribute name: StreamingFramePackingFormatsHTTP

Attribute definition: List of supported frame packing formats relevant for streaming of stereoscopic 3-D video over HTTP supported by the PSS application. The frame packing formats within scope for stereoscopic 3-D video include:

Frame Compatible Packing Formats: 1 = Vertical interleaving, 2 = Horizontal interleaving. 3 = Side-by-Side, 4 = Top-Bottom, 0 = Checkerboard

Full-Resolution per View Packing Formats: 5 = Temporal Interleaving

Component: Streaming

Type: Literal (Bag)

Legal values: List of integer values corresponding to the supported frame packing formats.

Resolution rule: Append

EXAMPLE: <StreamingFramePackingFormatsHTTP>
<rdf:Bag>
<rdf:li>3</rdf:li>
<rdf:li>4</rdf:li>
</rdf:Bag>
</StreamingFramePackingFormatsHTTP>

Attribute name: ThreeGPFramePackingFormats

Attribute definition: List of supported frame packing formats relevant for stereoscopic 3-D video that can be included in a 3GP file and handled by the PSS application.

Component: ThreeGPFileFormat

Type: Literal (Bag)

Legal values: List of integer values corresponding to the supported frame packing formats. Integer values shall be either 3 or 4 corresponding to the Side-by-Side and Top-and-Bottom frame packing formats respectively.

Resolution rule: Append

EXAMPLE: <ThreeGPFramePackingFormats>
<rdf:Bag>
<rdf:li>3</rdf:li>
<rdf:li>4</rdf:li>
</rdf:Bag>
</ThreeGPFramePackingFormats>

In some embodiments, a media presentation, as described in MPD, for example, may include attributes and elements common to Adaptation Set, Representation, and SubRepresentation. One such common element may be a FramePacking element. A FramePacking element may specify frame packing arrangement information of the video media component type. When no FramePacking element is provided for a video component, frame-packing may not be used for the video media component.

The FramePacking element may include an @schemeIdUri attribute that includes a uniform resource indicator (URI) to identify the frame packing configuration scheme employed. In some embodiments, the FramePacking element may further include an @value attribute to provide a value for the descriptor element.

In some embodiments, multiple FramePacking elements may be present. If so, each element may contain sufficient information to select or reject the described representation.

If the scheme or the value of all FramePacking elements are not recognized, the client may ignore the described Representations. A client may reject the Adaptation Set on the basis of observing a FramePacking element.

For Adaptation Sets or Representations that contain a video component that conforms to ISO/IEC Information technology -- Coding of audio-visual objects -- Part 10: Advanced Video Coding (ISO/IEC 14496-10:2012), a uniform resource number for FramePackin@schemeIdUri may be urn:mpeg:dash:14496:10:frame_packing_arrangement_type:2011, that may be defined to indicate the frame-packing arrangement as defined by Table D-8 of the ISO/IEC 14496-10:2012 ('Defintion of frame_packing_arrangement_type') to be contained in the FramePacking element. The @value may be the 'Value' column as specified in Table D-8 of the ISO/IEC 14496-10:2012 and may be interpreted according to the 'Interpretation' column in the same table.

Figure 5 illustrates a method 500 of signaling 3-D video device capabilities in accordance with some embodiments. Method 500 may be performed by components of a UE, e.g., UE 104. In some embodiments, the UE may include and/or have access to one or more computer-readable media having instructions stored thereon, that, when executed, cause the UE, or components thereof, to perform the method 500.

At 504, the UE may determine device capability information. As described above, the device capability information may include information as to the decoding and rendering capabilities of a media player. In some embodiments, a content management module, located on the UE or elsewhere, may determine this information by running one or more scripts on the UE to directly test the capabilities. In other embodiments, the content management module may access one or more stored files that contain the relevant information.

At 508, the UE may provide device capability information to the media server 116 or device profile server 144, including stereoscopic 3-D video decoding and rendering capabilities of the media player at the UE. As described above, the device capability information may include one or more format type attributes that represent a list of frame types supported by a streaming application of the UE. In some embodiments, the device capability information may be provided prior to or after the request at 512.

In some embodiments, some or all of the device capability information may be provided to the media server by another entity, e.g., a device profile server.

At 512, the UE may request 3-D video content. In some embodiments, the request may be in accordance with appropriate streaming/transport protocols, e.g., HTTP, RTP, RTSP, DASH, MBMS, PSS, IMS_PSS_MBMS, etc. The request may be directed to the media server and may include a uniform resource locator (URL) or some other indicator of the requested content or portions thereof. In some embodiments, the temporary adjustment to device capability information (e.g., via ProfDiff signaling) may also be provided along with the request at 508. Accordingly, the UE may supplement the profile information retrieved from the device profile server with extra attributes or overrides for attributes already defined in its device capability profile, based on ProfDiff signaling. In one example, such a temporary adjustment may be triggered by user preferences, for example if the user for a particular session only would like to receive two-dimensional (2-D) video even though the terminal is capable of rendering 3-D video.

At 516, the UE may receive the requested 3-D video content and render the content on a display of the UE. The rendering of the content may include a variety of processes such as, but not limited to, decoding, upconverting, unpacking, sequencing, etc.

Figure 6 illustrates a method 600 of signaling 3-D video content in accordance with some embodiments. Method 600 may be performed by components of a media server, e.g., media server 116. In some embodiments, the media server may include and/or have access to one or more computer-readable media having instructions stored thereon, that, when executed, cause the media server, or components thereof, to perform the method 600.

At 604, the media server may determine device capability information. In some embodiments, the media server may determine the device capability information by receiving, e.g., as part of session-level signaling, the information from the UE or a device profile server.

At 608, the media server may receive a request for 3-D video content. In some embodiments, the request may be in accordance with appropriate streaming/transport protocols, e.g., HTTP, RTP, RTSP, DASH, MBMS, PSS, IMS_PSS_MBMS, etc. The request may be from the UE and may include a universal resource locator (URL) or some other indicator of the requested content or portions thereof. In some embodiments, the request received at 608 may occur simultaneously with determination of the device capability information 604, before the determination, or after the determination. In some embodiments, the temporary adjustment to device capability information (e.g., via ProfDiff signaling) may also be received along with the request at 608. Accordingly, the media server may be supplemented with the profile information retrieved from the device profile server with extra attributes or overrides for attributes already defined in its device capability profile, based on ProfDiff signaling.

At 612, the media server may generate session description and/or metadata files to establish a streaming session, for example SDP file or a media presentation description (MPD) based on the device capability information accounting for the stereoscopic 3-D video decoding and rendering capabilities of the media player at the UE.

At 616, the media server may encode the 3-D video content in a format type indicated as being supported by the UE in the device capability information. The 3-D video content may then be streamed to the mobile device.

The components described herein, e.g., UE 104, media server 116, and/or device profile server 144, may be implemented into a system using any suitable hardware and/or software to configure as desired. Figure 7 illustrates, for one embodiment, an example system 700 comprising one or more processor(s) 704, system control logic 708 coupled with at least one of the processor(s) 704, system memory 712 coupled with system control logic

708, non-volatile memory (NVM)/storage 716 coupled with system control logic 708, a network interface 720 coupled with system control logic 708, and input/output (I/O) devices 732 coupled with system control logic 708.

The processor(s) 704 may include one or more single-core or multi-core processors. The processor(s) 704 may include any combination of general-purpose processors and dedicated processors (e.g., graphics processors, application processors, baseband processors, etc.).

System control logic 708 for one embodiment may include any suitable interface controllers to provide for any suitable interface to at least one of the processor(s) 704 and/or to any suitable device or component in communication with system control logic 708.

System control logic 708 for one embodiment may include one or more memory controller(s) to provide an interface to system memory 712. System memory 712 may be used to load and store data and/or instructions, e.g., logic 724. System memory 712 for one embodiment may include any suitable volatile memory, such as suitable dynamic random access memory (DRAM), for example.

NVM/storage 716 may include one or more tangible, non-transitory computer-readable media used to store data and/or instructions, e.g., logic 724. NVM/storage 716 may include any suitable non-volatile memory, such as flash memory, for example, and/or may include any suitable non-volatile storage device(s), such as one or more hard disk drive(s) (HDD(s)), one or more compact disk (CD) drive(s), and/or one or more digital versatile disk (DVD) drive(s), for example.

The NVM/storage 716 may include a storage resource physically part of a device on which the system 700 is installed or it may be accessible by, but not necessarily a part of, the device. For example, the NVM/storage 716 may be accessed over a network via the network interface 720 and/or over Input/Output (I/O) devices 732.

The logic 724, when executed by at least one of the processors 704 may cause the system to perform the operations described herein with respect to the UE 104, media server 116, and/or device profile server 144. The logic 724 may be disposed additionally/alternatively in other components of the system, e.g., in system control logic 708, and may include any combination of hardware, software, or firmware components.

Network interface 720 may have a transceiver 722 to provide a radio interface for system 700 to communicate over one or more network(s) and/or with any other suitable device. In various embodiments, the transceiver 722 may be integrated with other components of system 700. For example, the transceiver 722 may include a processor of the processor(s) 704, memory of the system memory 712, and NVM/Storage of NVM/Storage 716. Network interface 720 may include any suitable hardware and/or firmware. Network interface 720 may include a plurality of antennas to provide a multiple input, multiple output radio interface. Network interface 720 for one embodiment may include, for example, a wired network adapter, a wireless network adapter, a telephone modem, and/or a wireless modem.

For one embodiment, at least one of the processor(s) 704 may be packaged together with logic for one or more controller(s) of system control logic 708. For one embodiment, at least one of the processor(s) 704 may be packaged together with logic for one or more controllers of system control logic 708 to form a System in Package (SiP). For one embodiment, at least one of the processor(s) 704 may be integrated on the same die with logic for one or more controller(s) of system control logic 708. For one embodiment, at least one of the processor(s) 704 may be integrated on the same die with logic for one or more controller(s) of system control logic 708 to form a System on Chip (SoC).

In various embodiments, the I/O devices 732 may include user interfaces designed to enable user interaction with the system 700, peripheral component interfaces designed to enable peripheral component interaction with the system 700, and/or sensors designed to determine environmental conditions and/or location information related to the system 700.

In various embodiments, the user interfaces could include, but are not limited to, a display for rendering 3-D video (e.g., a liquid crystal display, a touch screen display, an auto-stereoscopic display, etc.), a speaker, a microphone, one or more cameras (e.g., a still camera and/or a video camera), a flashlight (e.g., a light emitting diode flash), and a keyboard.

In various embodiments, the peripheral component interfaces may include, but are not limited to, a non-volatile memory port, a universal serial bus (USB) port, an audio jack, and a power supply interface.

In various embodiments, the sensors may include, but are not limited to, a gyro sensor, an accelerometer, a proximity sensor, an ambient light sensor, and a positioning unit. The positioning unit may also be part of, or interact with, the network interface 720 to communicate with components of a positioning network, e.g., a global positioning system (GPS) satellite.

In various embodiments, the system 700 may be a mobile computing device such as, but not limited to, a laptop computing device, a tablet computing device, a netbook, a smartphone, etc. In various embodiments, system 700 may have more or less components, and/or different architectures.

Although certain embodiments have been illustrated and described herein for purposes of description, a wide variety of alternate and/or equivalent embodiments or implementations calculated to achieve the same purposes may be substituted for the embodiments shown and described without departing from the scope of the present disclosure. This application is intended to cover any adaptations or variations of the embodiments discussed herein. Therefore, it is manifestly intended that embodiments described herein be limited only by the claims and the equivalents thereof.

Claims

What is claimed is:

1. One or more computer-readable media having instructions that, when executed, cause a device to:

obtain, from a network entity, a streaming frame packing format attribute that includes a list of frame packing formats, supported by a client terminal of a wireless communication network, relevant for streaming of stereoscopic 3-D video over a transport protocol supported by a packet-switched streaming service (PSS) application on the client terminal, wherein the transport protocol is a real-time transport protocol (RTP) or a hypertext transfer protocol (HTTP);

adapt content based on the streaming frame packing format attribute;

generate a session description or metadata file to establish a streaming session based on the streaming frame packing format attribute; and

transmit the adapted content and the generated session description or metadata file to the client terminal.

2. The one or more computer-readable media of claim 1, wherein the list of frame packing formats includes an indication of a vertical interleaving frame compatible packing format, a horizontal interleaving frame compatible packing format, a side-by-side frame compatible packing format, a top-bottom frame compatible packing format, a checkerboard frame compatible packing format, or a temporal interleaving full-resolution per view packing format.

3. The one or more computer-readable media of either claim 1 or 2, wherein the list comprises a list of one or more integer values that respectively correspond to one or more supported frame packing formats.

4. The one or more computer-readable media of claim 3, wherein the list of one or more integer values includes a 1 to correspond to a vertical interleaving frame compatible packing format, a 2 to correspond to a horizontal interleaving frame compatible packing format, a 3 to correspond to a side-by-side frame compatible packing format, a 4 to correspond to a top-bottom frame compatible packing format, a 0 to correspond to a checkerboard frame compatible packing format, or a 5 to correspond to a temporal interleaving full-resolution per view packing format.

5. The one or more computer-readable media of either claim 1 or 2, wherein the transport protocol comprises an RTP.

6. The one or more computer-readable media of either claim 1 or 2, wherein the transport protocol comprises an HTTP.

7. The one or more computer-readable media of either claim 1 or 2, wherein the session description or metadata file is a real-time streaming protocol (RTSP) session description protocol (SDP) file or a dynamic adaptive streaming over hypertext transport protocol (DASH) media presentation description (MPD) file.

8. The one or more computer-readable media of either claim 1 or 2, wherein content is 3-D video content and the instructions, when executed, further cause the device to:
transcode the 3-D video content or convert a format of the 3-D video content based on the streaming frame packing format attribute.

9. The one or more computer-readable media of either claim 1 or 2, wherein the network entity is a device profile server or a user equipment that comprises the client terminal.

10. One or more computer-readable media having instructions that, when executed, cause a device to:
obtain, from a network entity, a frame packing format attribute that includes a list of one or more frame packing formats, supported by a user equipment, relevant for stereoscopic 3-D video that can be included in a 3rd Generation Partnership Project file format (3GP) file and handled by a packet-switched streaming service (PSS) application on the user equipment; and
transmit content to the user equipment based on the frame packing format attribute.

11. The one or more computer-readable media of claim 10, wherein the list of frame packing formats includes an indication of a side-by-side frame packing format or a top-bottom frame packing format.

12. The one or more computer-readable media of either claim 10 or 11, wherein the list comprises a list of one or more integer values that respectively correspond to one or more frame packing formats.

13. The one or more computer-readable media of claim 12, wherein the list of one or more integer values includes a 3 to correspond to side-by-side frame packing format or a 4 to correspond to a top-bottom frame packing format.

14. An apparatus comprising:
a media player to decode and render stereoscopic three-dimensional (3-D) video content, wirelessly received by a mobile device, on a display of the mobile device; and
a content management module to:
determine device capability information including 3-D video codecs and formats supported by the media player;

transmit one or more messages to a media server or a device profile server, the one or more messages including the device capability information;
transmit at least one message to a media server, the at least one message including a request for stereoscopic 3-D video content and any temporary adjustments on the device capability information.

15. The apparatus of claim 14, wherein the device capability information has a 3-D format attribute that includes a format type supported by the media player circuitry.

16. The apparatus of claim 15, wherein the format type is a frame-packing format type that corresponds to a frame compatible packing format or a full-resolution per view packing format, a simulcast format type, or a two-dimensional plus auxiliary format type.

17. The apparatus of claim 16, wherein the format type is a frame-packing format type that is a frame compatible packing format with a value to indicate vertical interleaving, horizontal interleaving, side-by-side, top-bottom, or checkerboard.

18. The apparatus of any of claims 14-17, wherein the 3-D video content is wirelessly received by the mobile device via a packet-switched streaming service.

19. The apparatus of any of claims 14-17, wherein the content management module is to transmit device capability information in a session initiation protocol (SIP) SUBSCRIBE message to an Internet Protocol Multimedia Core Network subsystem during service discovery.

20. The apparatus of claim 19, wherein, subsequent to service discovery, the content management module is to use SIP signaling to update set of supported 3-D video codecs and formats.

21. The apparatus of any of claims 14-17, wherein the media player is further configured to:
receive a stream including the requested stereoscopic 3-D video content, wherein the media player is further configured to receive the stream according to a dynamic adaptive streaming over hypertext transport portocol (DASH) protocol; a packet-switched streaming (PSS) protocol; or an Internet protocol multimedia subsystem (IMS)-based PSS and multimedia broadcast/multicast (MBMS) services protocol.

22. The apparatus of claim 21, wherein the media player is further configured to receive a session description protocol (SDP) file or a media presentation description (MPD) metadata file associated with the stream.

23. A mobile device comprising the apparatus of any of claims 14-17.

24. The mobile device of claim 23, further comprising:
an auto-stereoscopic display to render the 3-D video content under control of the media player.

25. One or more computer-readable media having instructions that, when executed, cause a device to:
obtain a request for 3-D capability information related to a user equipment; and
provide a streaming frame packing format attribute that includes a list of frame packing formats, supported by the user equipment, relevant for streaming of stereoscopic 3-D video over a transport protocol supported by a packet-switched streaming service (PSS) application on the user equipment, wherein the transport protocol is a real-time transport protocol (RTP) or a hypertext transfer protocol (HTTP).

26. The one or more computer-readable media of claim 25, wherein the list of frame packing formats includes an indication of a vertical interleaving frame compatible packing format, a horizontal interleaving frame compatible packing format, a side-by-side frame compatible packing format, a top-bottom frame compatible packing format, a checkerboard frame compatible packing format, or a temporal interleaving full-resolution per view packing format.

27. The one or more computer-readable media of claim 25, wherein the transport protocol comprises an RTP.

28. The one or more computer-readable media of claim 25, wherein the transport protocol comprises an HTTP.

29. The one or more computer-readable media of any of claims 25-28, wherein the list comprises a list of one or more integer values that respectively correspond to one or more supported frame packing formats.

30. The one or more computer-readable media of claim 29, wherein the list of one or more integer values includes a 1 to correspond to a vertical interleaving frame compatible packing format, a 2 to correspond to a horizontal interleaving frame compatible packing format, a 3 to correspond to a side-by-side frame compatible packing format, a 4 to correspond to a top-bottom frame compatible packing format, a 0 to correspond to a checkerboard frame compatible packing format, or a 5 to correspond to a temporal interleaving full-resolution per view packing format.

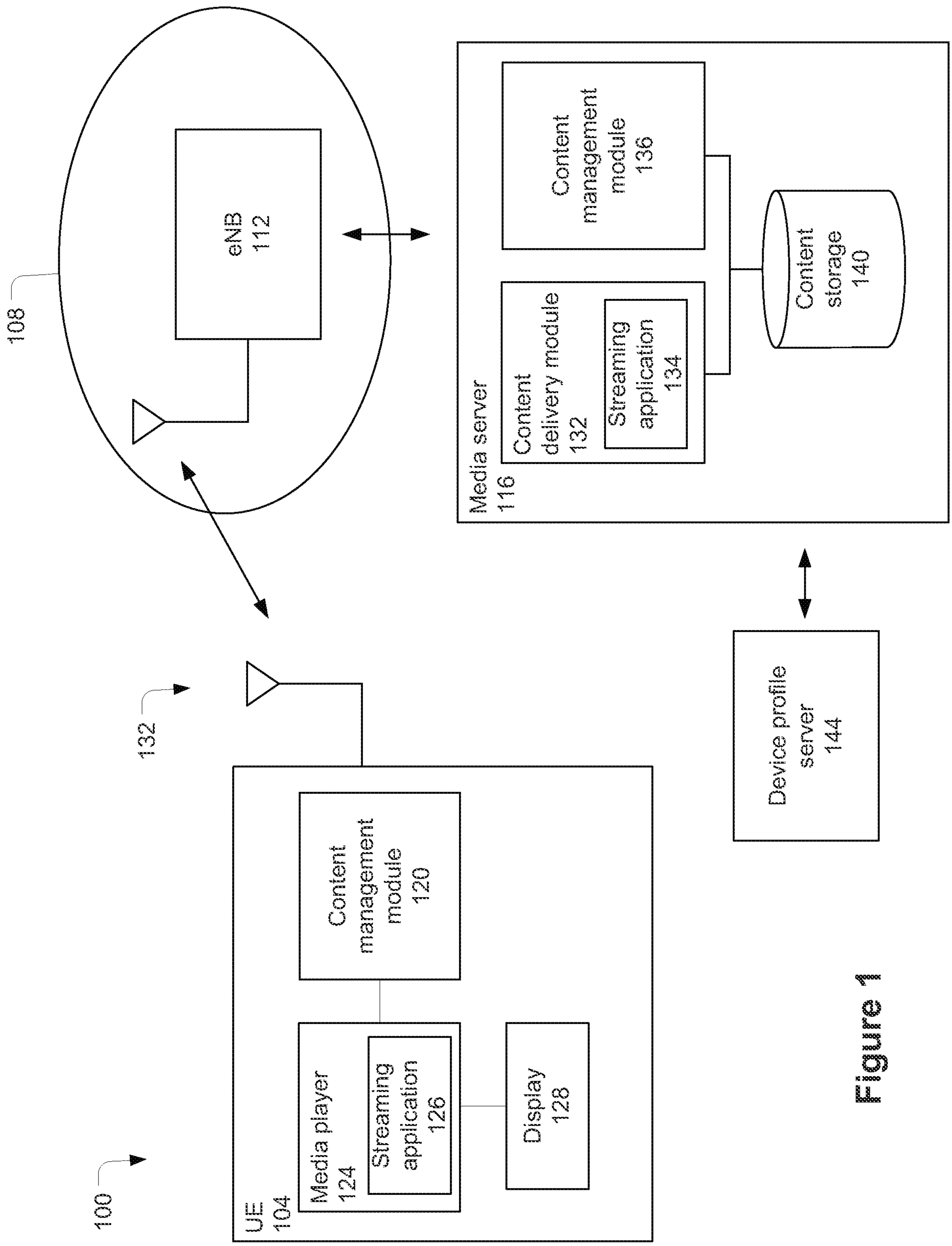
100
UE
104
Media player
124
Streaming
application
126
Display
128
Content
management
module
120
132
108
eNB
112
Media server
116
Content
delivery module
132
Streaming
application
134
Content
management
module
136
Content
storage
140
Device profile
server
144

Figure 1

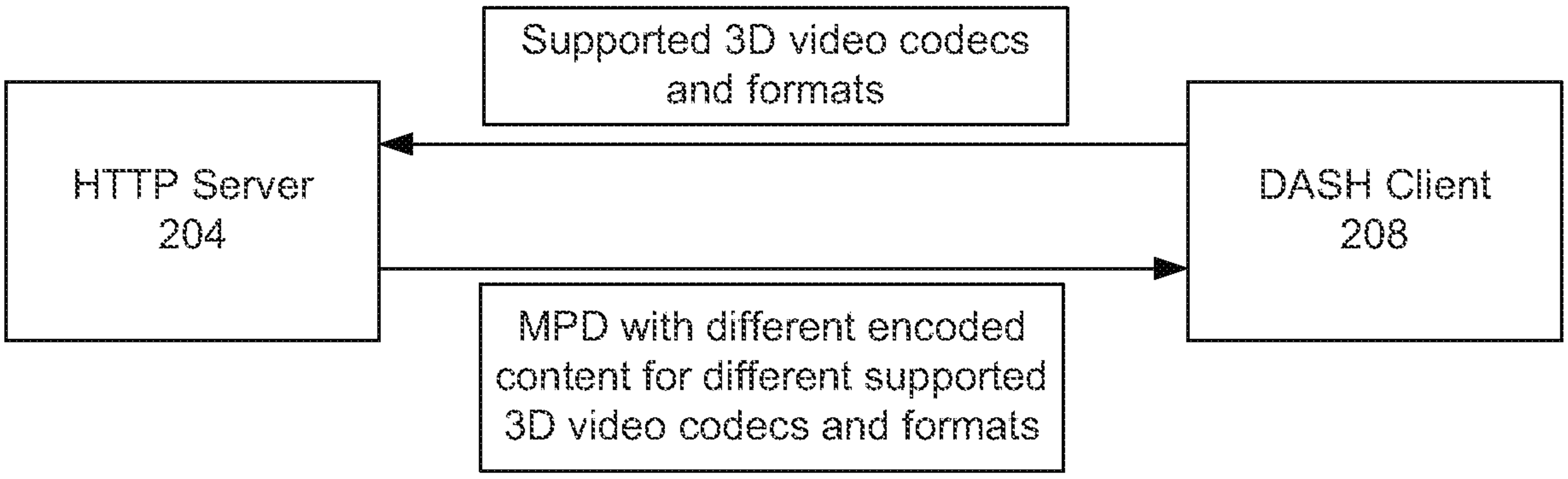

Figure 2a

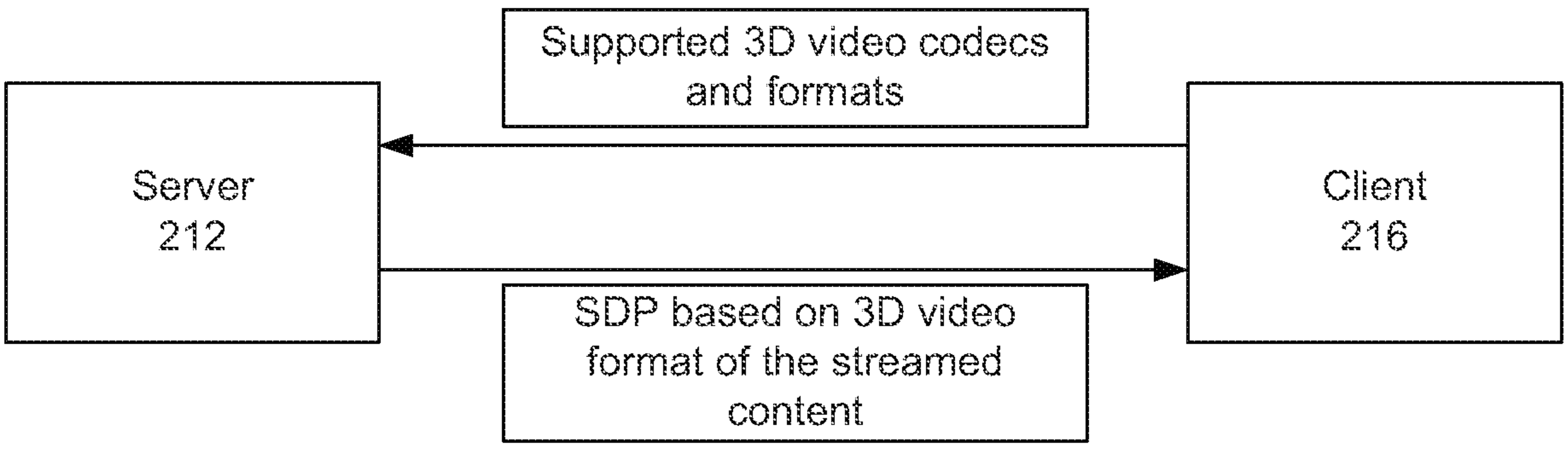

Figure 2b

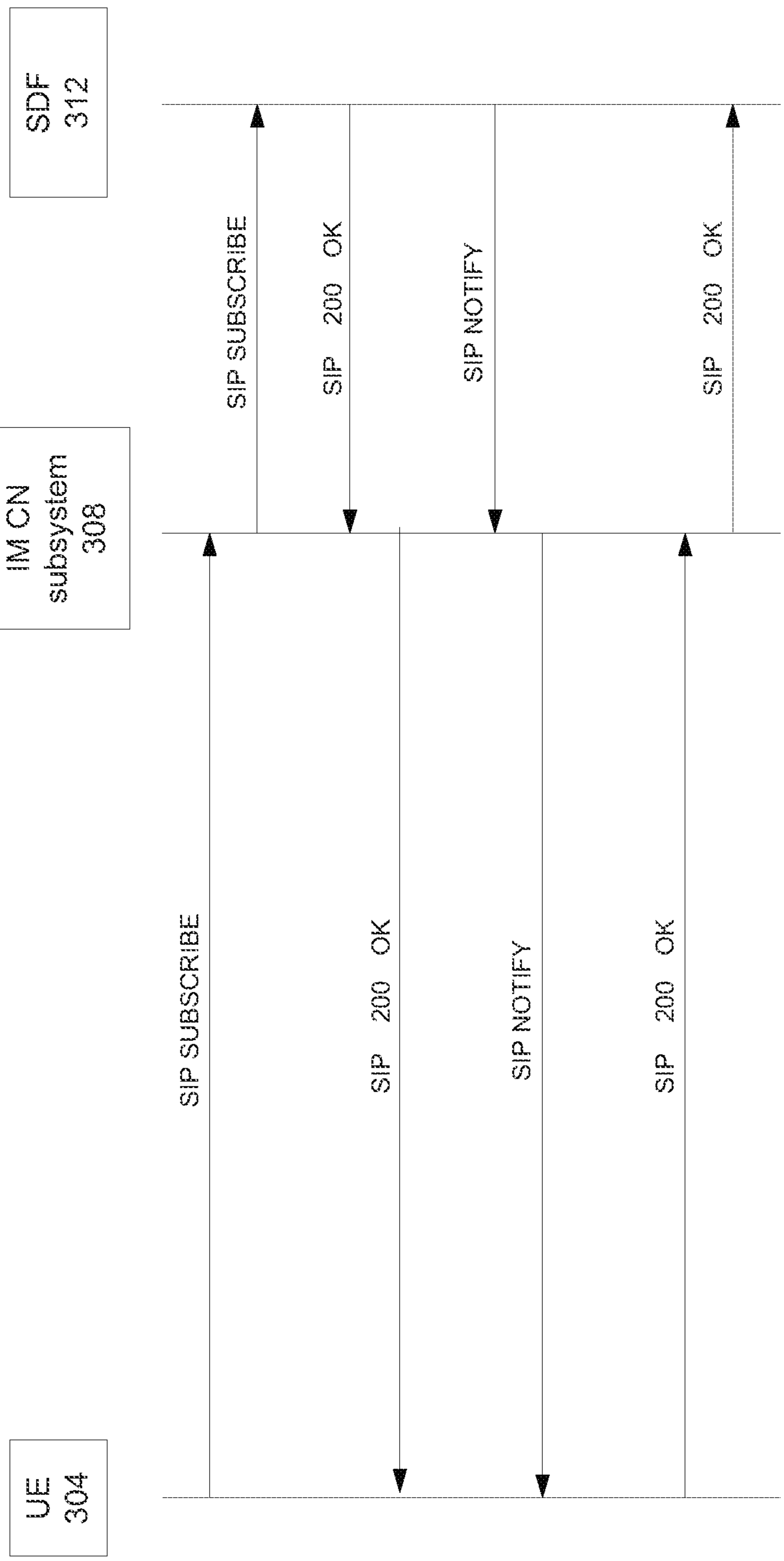
SDF
312
IM CN
subsystem
308
UE
304
SIP SUBSCRIBE
SIP 200 OK
SIP NOTIFY
SIP 200 OK
SIP SUBSCRIBE
SIP 200 OK
SIP NOTIFY
SIP 200 OK

Figure 3

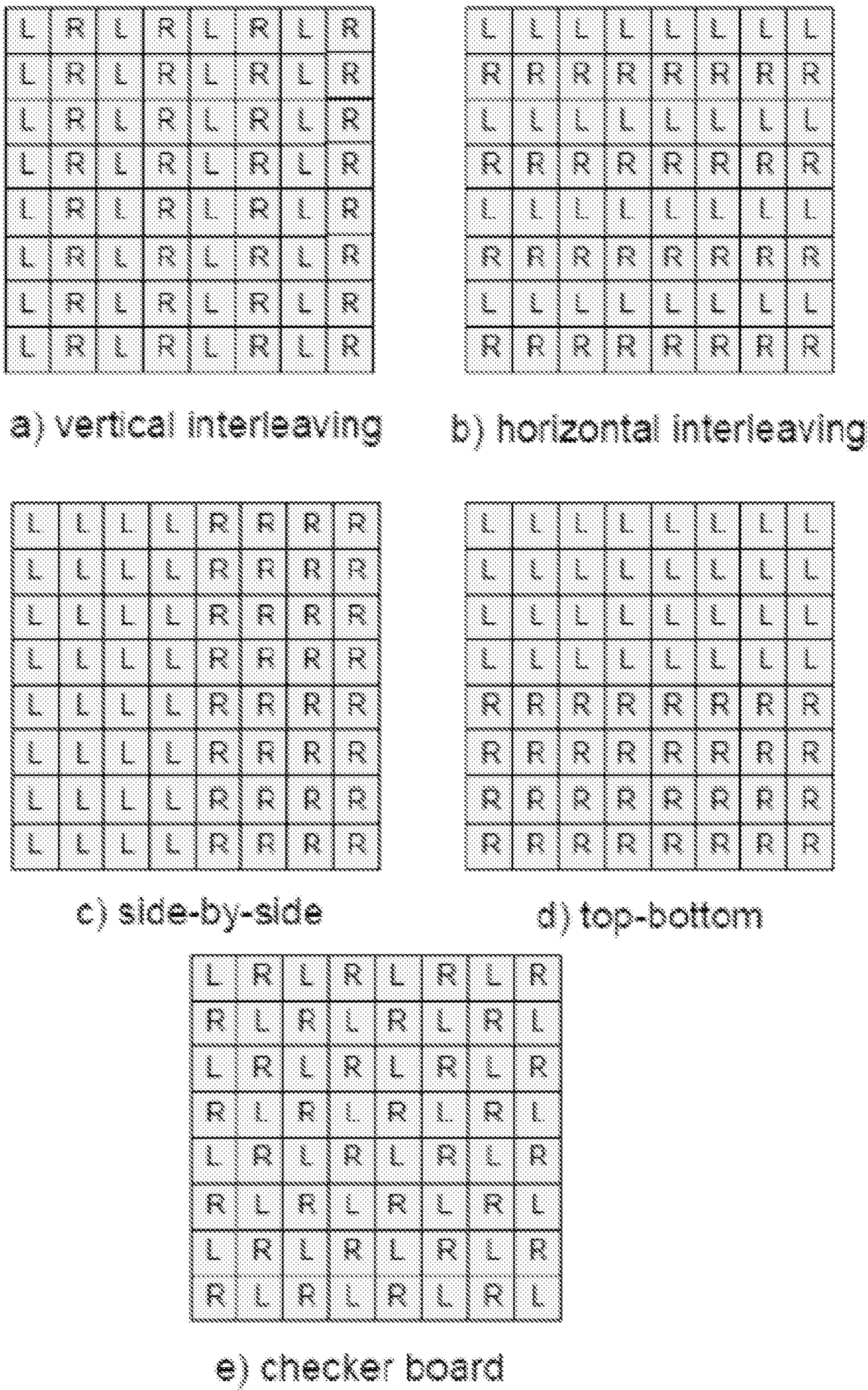
a) vertical interleaving
b) horizontal interleaving
c) side-by-side
d) top-bottom
e) checker board

Figure 4

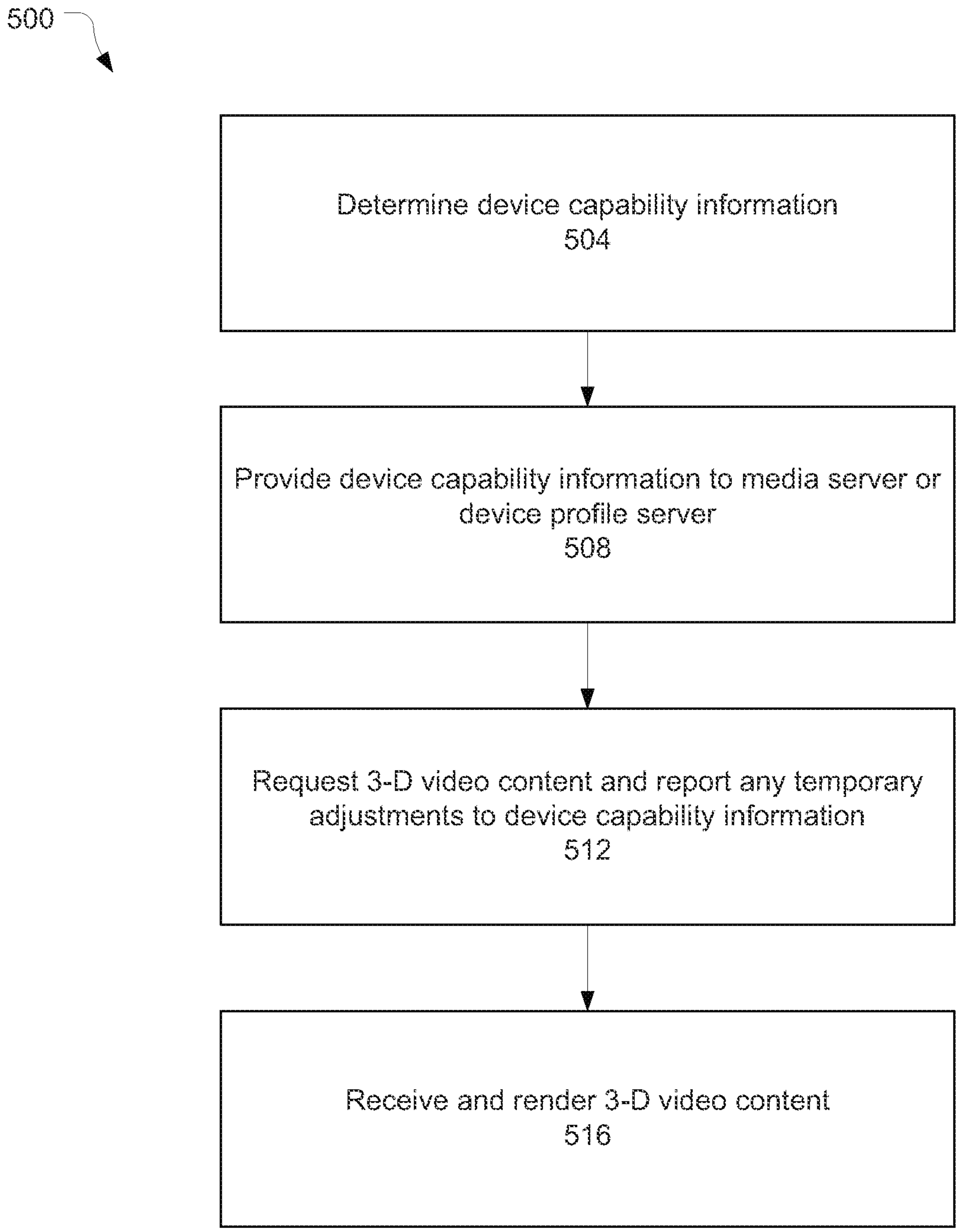
500
Determine device capability information
504
Provide device capability information to media server or device profile server
508
Request 3-D video content and report any temporary adjustments to device capability information
512
Receive and render 3-D video content
516

Figure 5

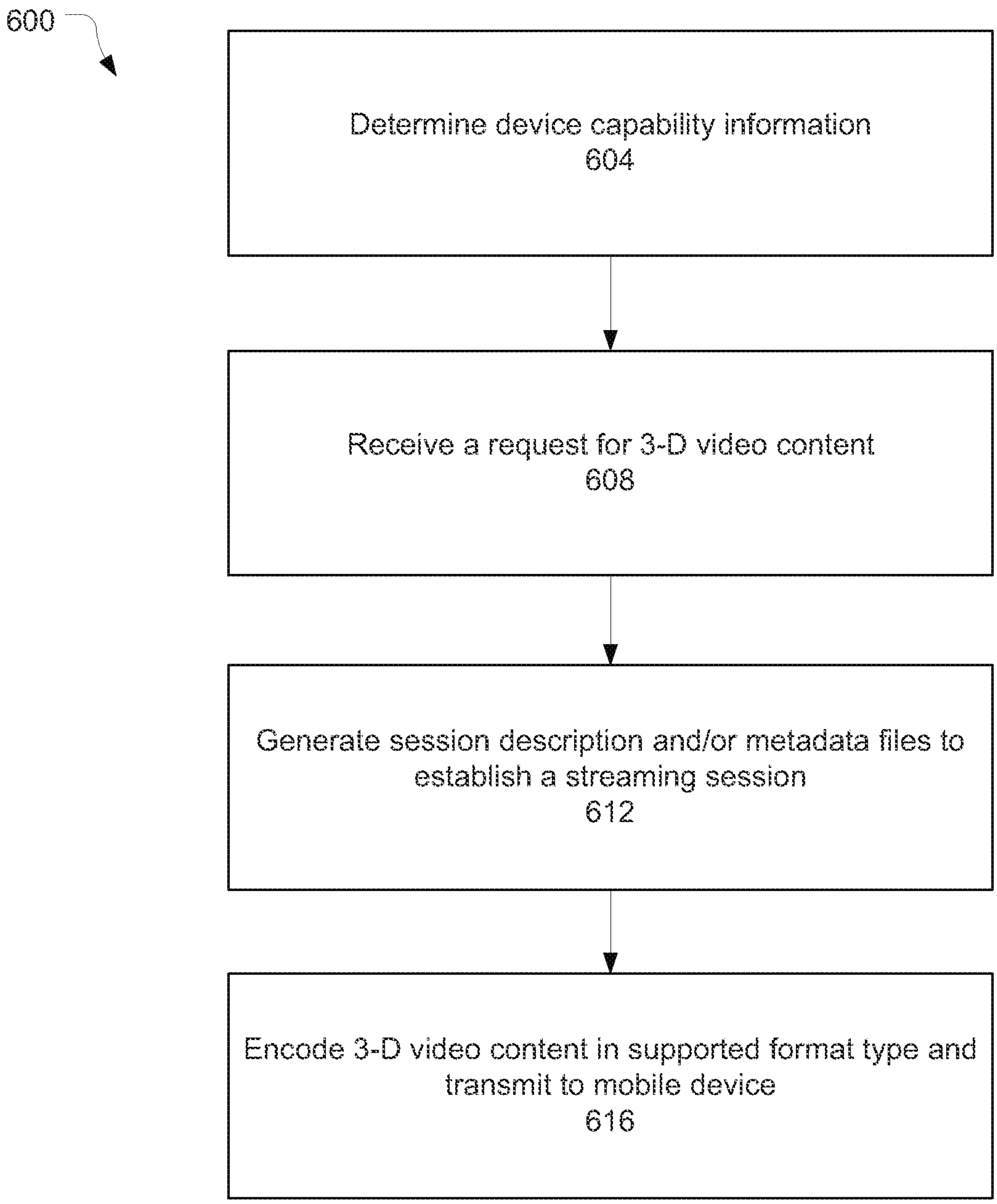
600
Determine device capability information
604
Receive a request for 3-D video content
608
Generate session description and/or metadata files to establish a streaming session
612
Encode 3-D video content in supported format type and transmit to mobile device
616

Figure 6

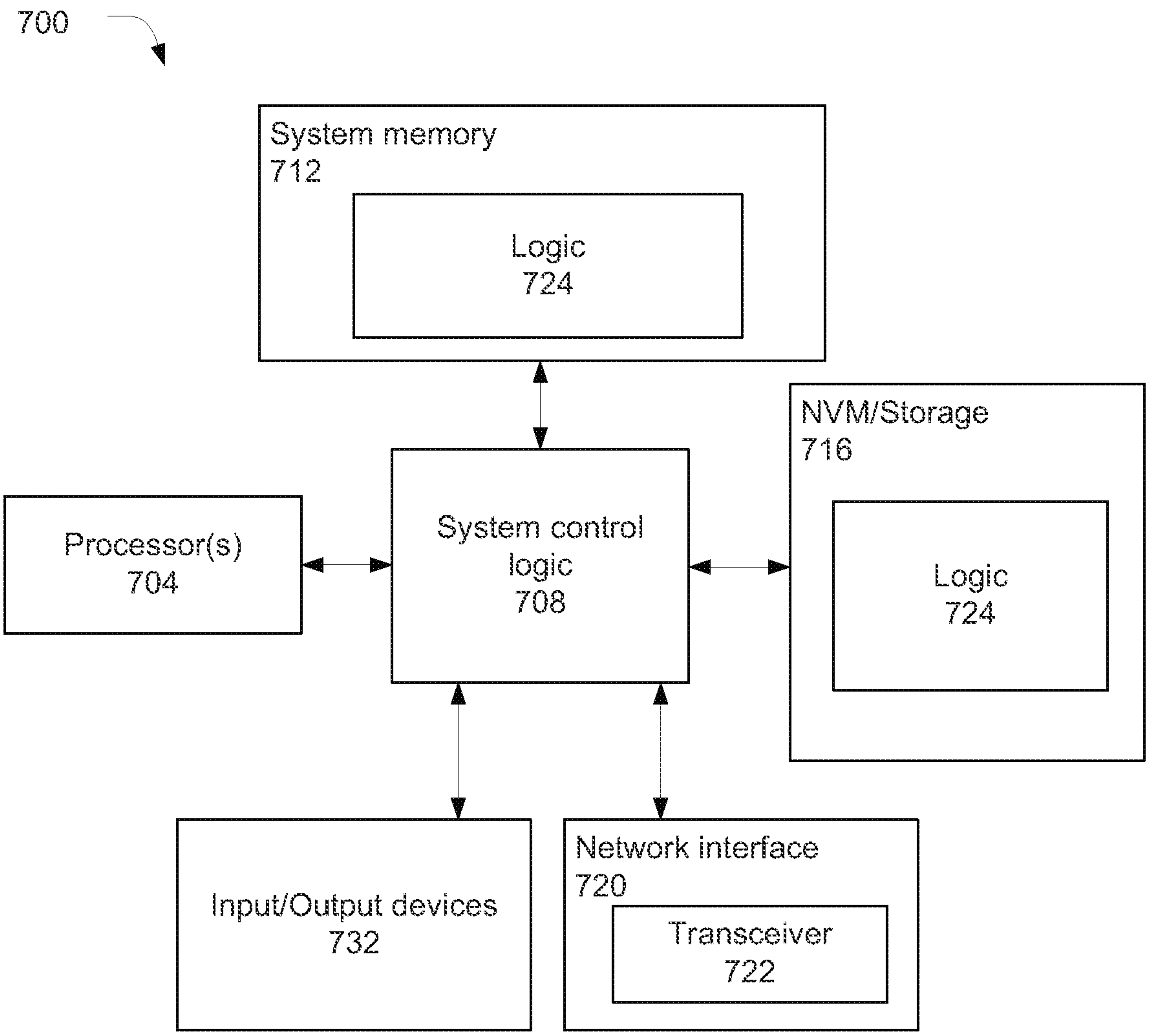
700
System memory
712
Logic
724
NVM/Storage
716
Processor(s)
704
System control
logic
708
Logic
724
Input/Output devices
732
Network interface
720
Transceiver
722

Figure 7

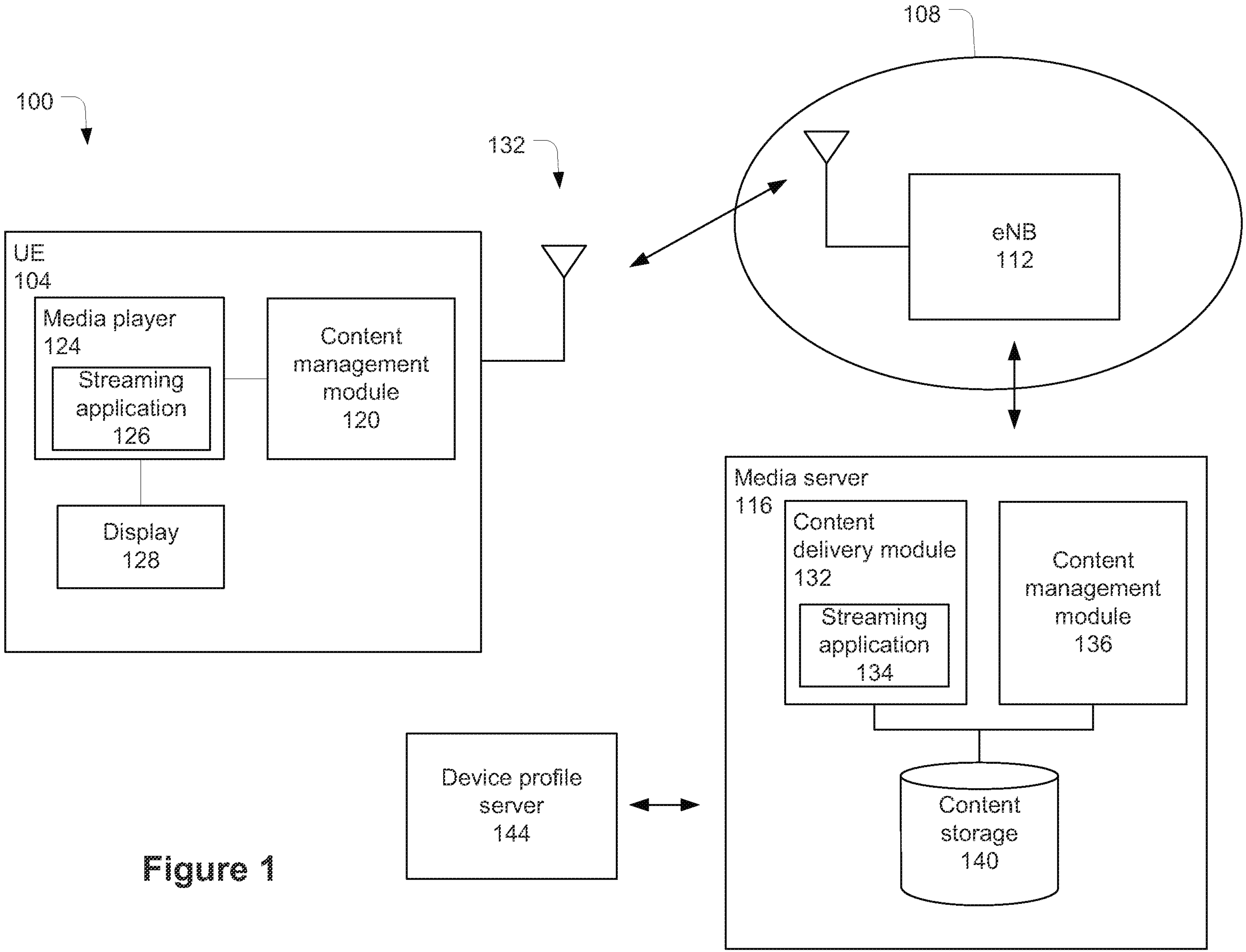
100
108
132
UE
104
Media player
124
Streaming application
126
Content management module
120
Display
128
eNB
112
Media server
116
Content delivery module
132
Streaming application
134
Content management module
136
Content storage
140
Device profile server
144

Figure 1