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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
01 November 2018 (01.11.2018)

WIPO | PCT

(10) International Publication Number
WO 2018/199987 A1

(51) International Patent Classification:
H04N 21/439 (2011.01) *H04H 60/56* (2008.01)
H04N 21/434 (2011.01)

(21) International Application Number:
PCT/US2017/030153

(22) International Filing Date:
28 April 2017 (28.04.2017)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant: **HEWLETT-PACKARD DEVELOPMENT COMPANY, L.P.** [US/US]; 11445 Compaq Center Drive W., Houston, TX 77070 (US).

(72) Inventors: **BHARITKAR, Sunil**; 1501 Page Mill Road, Palo Alto, CA 94304 (US). **LU, Maureen, Min-Chaun**; 10F, No.66, Jing Mao 2 Rd., NanGang District, 11568 Taipei (TW).

(74) Agent: **BURROWS, Sarah, E.** et al.; HP INC., 3390 E. Harmony Road, Mail Stop 35, Fort Collins, CO 80528 (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, DJ, 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, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, 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, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, 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,

(54) Title: AUDIO TUNING PRESETS SELECTION

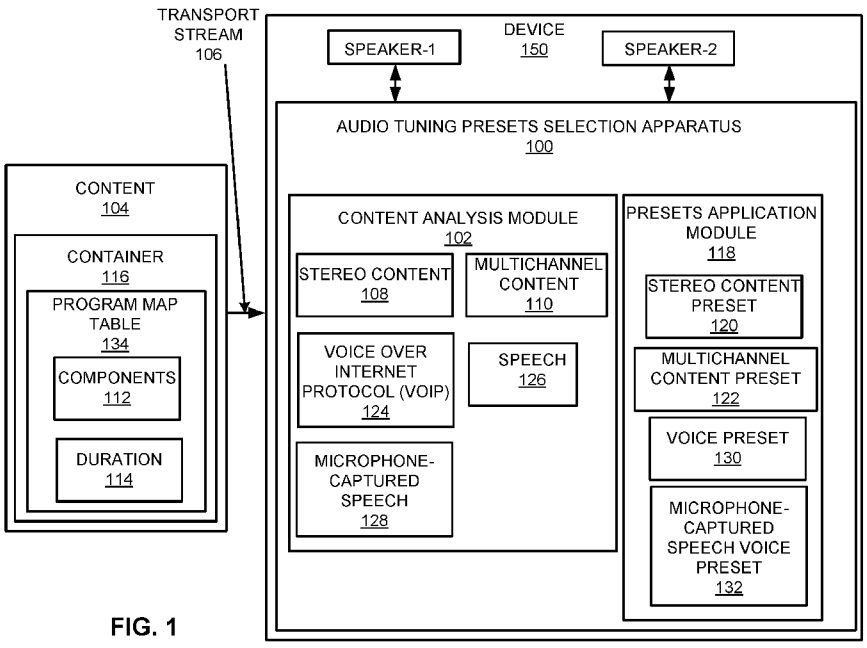


FIG. 1

(57) Abstract: In some examples, audio tuning presets selection may include determining whether content included in a transport stream includes stereo content or multichannel content by analyzing components of the content included in the transport stream, and a duration of the content included in the transport stream. In response to a determination that the content included in the transport stream includes the stereo content, a stereo content preset may be applied to the content included in the transport stream. Alternatively, in response to a determination that the content included in the transport stream includes the multichannel content, a multichannel content preset may be applied to the content included in the transport stream.

WO 2018/199987 A1

[Continued on next page]

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).

Declarations under Rule 4.17:

— *as to the identity of the inventor (Rule 4.17(i))*

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

Published:

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

AUDIO TUNING PRESETS SELECTION

BACKGROUND

[0001] Devices such as notebooks, desktop computers, mobile telephones, tablets, and other such devices may include speakers or utilize headphones to reproduce sound. The sound emitted from such devices may be subject to various processes that modify the sound quality.

BRIEF DESCRIPTION OF DRAWINGS

[0002] Features of the present disclosure are illustrated by way of example and not limited in the following figure(s), in which like numerals indicate like elements, in which:

[0003] Figure 1 illustrates an example layout of an audio tuning presets selection apparatus;

[0004] Figure 2 illustrates an example graphical user interface of device consumer selectable presets;

[0005] Figure 3 illustrates an example layout of a Moving Picture Experts Group (MPEG)-2 transport stream for terrestrial and satellite;

[0006] Figure 4 illustrates an example layout of an MPEG2 transport stream for Internet Protocol television;

[0007] Figure 5 illustrates an example MPEG2 transport stream for Internet Protocol television;

[0008] Figure 6 illustrates an example of MPEG2 transport stream multiplexing video, audio, and program information via a Program Map table;

[0009] Figure 7 illustrates an example MPEG2 transport stream Program Map table;

[0010] Figure 8 illustrates an example of MPEG2 transport stream Advanced Audio Coding-Audio Data Transport Stream;

[0011] Figure 9 illustrates an example MPEG2 transport stream Program Map table with audio stream types identifying cinematic content;

[0012] Figure 10 illustrates an example content classifier;

[0013] Figure 11 illustrates an example of audio decoder element values indicating number of channels;

[0014] Figure 12 illustrates an example of MPEG Advanced Audio Coding

based bitstream syntax;

[0015] Figure 13 illustrates an example of MPEG Advanced Audio Coding based channel configurations;

[0016] Figure 14 illustrates an example of MP-4 metadata;

[0017] Figure 15 illustrates an example of average duration of movie content along with some specific movies;

[0018] Figure 16 illustrates an example distribution of a video length;

[0019] Figure 17 illustrates examples of distribution of various genres in seconds;

[0020] Figure 18 illustrates an example block diagram for audio tuning presets selection;

[0021] Figure 19 illustrates an example flowchart of a method for audio tuning presets selection; and

[0022] Figure 20 illustrates a further example block diagram for audio tuning presets selection.

DETAILED DESCRIPTION

[0023] For simplicity and illustrative purposes, the present disclosure is described by referring mainly to examples. In the following description, numerous specific details are set forth in order to provide a thorough understanding of the present disclosure. It will be readily apparent however, that the present disclosure may be practiced without limitation to these specific details. In other instances, some methods and structures have not been described in detail so as not to unnecessarily obscure the present disclosure.

[0024] Throughout the present disclosure, the terms "a" and "an" are intended to denote at least one of a particular element. As used herein, the term "includes" means includes but not limited to, the term "including" means including but not limited to. The term "based on" means based at least in part on.

[0025] Audio tuning presets selection apparatuses, methods for audio tuning presets selection, and non-transitory computer readable media having stored thereon machine readable instructions to provide audio tuning presets selection are disclosed herein. The apparatuses, methods, and non-transitory computer readable media disclosed herein provide for identification of stereo or multichannel content, as well as differentiation between stereo music and downmixed cinematic content (e.g., downmixed from 5.1 to stereo). In this regard, cinematic content may be multichannel (e.g., 5.1, 7.1, etc., where 5.1 represents “five point one” and includes a six channel surround sound audio system, 7.1 represents “seven point one” and includes an eight channel surround sound audio system, etc.). The identification of stereo or multichannel content may provide for the correct preset to be applied depending on the type of content, without the need for consumer intervention. Additionally, a separate preset may be applied to enhance speech or dialog clarity based on detection of voice in Voice over Internet Protocol (VoIP), or voice in cinematic content.

[0026] With respect to audio tuning, personal devices including loudspeakers may need to be tuned or calibrated in order to reduce the effects of loudspeaker

and/or room acoustics, while also maximizing the quality of experience (QoE) for content involving audio or speech. Depending on the type of content being listened to, the tuning (viz., the type of preset and the corresponding value of the preset) may need to be applied correctly. For example, with music being stereo (e.g., 2-channels), a device may allow for bass, mid-range, and treble frequency control presets depending on device capability. In the case of cinematic content, which may be, for example, 5.1 channels (or next-generation object-based), it is technically challenging to determine a different set of presets to control various elements of the cinematic mix reproduced on personal devices. For example, a device may include a preset type that is the same for both music and cinematic/movie content, as if both were stereo audio content, whereas the actual values assigned to those presets may be different.

[0027] In devices such as all-in-one computers, desktops, etc., an interface may be provided for a consumer to select one of three pre-programmed (i.e., tuned) presets from movie, music, and voice. Figure 2 illustrates an example graphical user interface of device consumer selectable presets. Referring to Figure 2, the graphical user interface may include XYZ company tuning that is implemented, for example, as a Windows Audio processing object (APO), and may not be readily discernible to a consumer. Even if the consumer is aware, the consumer may need to manually apply the correct preset each time to the appropriate content. This process may be error-prone due to factors such as consumer unawareness, whether the consumer remembers to apply the correct preset, application of the wrong preset by the consumer to a particular type of content, etc. These aspects may degrade the desired quality of experience.

[0028] In order to address at least these technical challenges associated with determination of a different set of presets to control various elements of a cinematic mix reproduced on personal devices, and application of a correct preset to appropriate content, the apparatuses, methods, and non-transitory computer readable media disclosed herein provide for identification of stereo or multichannel

content, as well as differentiation between stereo music (with or without video) and stereo downmixed cinematic content (downmixed from 5.1 to stereo). This identification provides for the correct preset to be applied depending on the type of content, without the need for consumer intervention, at a reference playback level. Further, based on the detection of voice in VoIP (or cinematic content where voice is in the center channel), a voice preset may be applied to enhance speech or dialog clarity. Yet further, a modified voice preset may be applied to microphone-captured speech based on detection of a keyword, where a preset may be used for enhancing the speech formant peaks and widths and adding specific equalization, compression, speech-rate change, etc..

[0029] For the apparatuses, methods, and non-transitory computer readable media disclosed herein, modules, as described herein, may be any combination of hardware and programming to implement the functionalities of the respective modules. In some examples described herein, the combinations of hardware and programming may be implemented in a number of different ways. For example, the programming for the modules may be processor executable instructions stored on a non-transitory machine-readable storage medium and the hardware for the modules may include a processing resource to execute those instructions. In these examples, a computing device implementing such modules may include the machine-readable storage medium storing the instructions and the processing resource to execute the instructions, or the machine-readable storage medium may be separately stored and accessible by the computing device and the processing resource. In some examples, some modules may be implemented in circuitry.

[0030] Figure 1 illustrates an example layout of an audio tuning presets selection apparatus (hereinafter also referred to as “apparatus 100”).

[0031] In some examples, the apparatus 100 may include or be provided as a component of a device such as a notebook, a desktop computer, a mobile telephone, a tablet, and other such devices. For the example of Figure 1, the apparatus 100 is illustrated as being provided as a component of a device 150,

which may include a notebook, a desktop computer, a mobile telephone, a tablet, and other such devices.

[0032] Referring to Figure 1, the apparatus 100 may include a content analysis module 102 to determine whether content 104 included in a transport stream 106 includes stereo content 108 or multichannel content 110 by analyzing components 112 of the content 104 included in the transport stream 106, and a duration 114 of the content 104 included in the transport stream 106. The components 112 of the content 104 included in the transport stream may be included in a container 116 included in the transport stream 106. For example, the presentation timestamp (PTS) metadata field may be extracted at the beginning and at the end of the content, and the difference in the timestamps may be used to determine the duration of the content. The presentation timestamp may represent the exact time that a frame needs to display. The presentation timestamp may be determined from streams, such as MPEG-2, MPEG-4 or H.264 streams. Video, audio, and data in the same program stream may use the same base time. Therefore, the synchronization between channels in the same program stream may be achieved by using this time stamp.

[0033] In response to a determination that the content 104 included in the transport stream includes the stereo content 108, a presets application module 118 is to apply a stereo content preset 120 to the content 104 included in the transport stream 106. Alternatively, in response to a determination that the content 104 included in the transport stream 106 includes the multichannel content 110, the presets application module 118 is to apply a multichannel content preset 122 to the content 104 included in the transport stream 106.

[0034] According to an example, the content analysis module 102 is to determine whether the stereo content 108 includes a video, whether the stereo content 108 does not include the video, and whether the stereo content 108 is downmixed cinematic content. In response to a determination that the stereo content 108 includes the video, that the stereo content 108 does not include the

video, or that the stereo content 108 is downmixed cinematic content, the presets application module 118 is to apply a corresponding type (e.g., from types related to video, no video, and downmixed cinematic content) of the stereo content preset 120 to the content 104 included in the transport stream 106.

[0035] According to an example, the content analysis module 102 is to determine whether the content 104 included in the transport stream 106 includes voice in Voice over Internet Protocol (VoIP) 124, or speech 126. In response to a determination that the content 104 included in the transport stream 106 includes the voice in VoIP 124, or the speech 126, the presets application module 118 is to apply a voice preset 130 to the content 104 included in the transport stream 106.

[0036] According to an example, the content analysis module 102 is to determine whether the content 104 included in the transport stream 106 includes microphone-captured speech 128. In response to a determination that the content 104 included in the transport stream 106 includes the microphone-captured speech 128, the presets application module 118 is to apply a microphone-captured speech voice preset 132 to the content 104 included in the transport stream 106.

[0037] According to an example, the content analysis module 102 is to determine, based on the analysis of the components 112 in a Program Map table 134 included in the container 116, whether the content 104 included in the transport stream 106 includes audio content, or audio and video content. The components 112 may include audio frames for the audio content. Further, in response to a determination that the content 104 included in the transport stream 106 includes the audio content, or the audio and video content, the presets application module 118 is to selectively apply the stereo content preset 120 or the multichannel content preset 122 to the content 104 included in the transport stream 106.

[0038] According to an example, the content analysis module 102 is to determine, based on the analysis of the components 112 in the Program Map table 134 included in the container 116, whether the content 104 included in the transport stream 106 includes audio content, or audio-for-video content. In response to a

determination that the content 104 included in the transport stream 106 includes the audio content, or the audio-for-video content, the presets application module 118 is to selectively apply the stereo content preset 120 or the multichannel content preset 122 to the content 104 included in the transport stream 106.

[0039] According to an example, the content analysis module 102 is to determine the duration 114 of the content 104 included in the transport stream 106 by analyzing a file-size and a data rate associated with the content 104 included in the transport stream 106. In this regard, the data rate may include a constant bitrate or a variable bitrate. Further, the content analysis module 102 is to analyze the duration 114 of the content 104 included in the transport stream 106 by comparing the duration 114 to predetermined durations for different types of stereo content and multichannel content.

[0040] With respect to detection of content type, the content analysis module 102 may rely on audio decoder, video decoder, transport stream and/or container file-format being used to extract or decode audio (e.g., no video) or audio/video content.

[0041] Figure 3 illustrates an example layout of a Moving Picture Experts Group (MPEG)-2 transport stream for terrestrial and satellite.

[0042] Referring to Figure 3, Figure 3 includes an example of how various signals may be transported using a transport stream 300, for example, using terrestrial/satellite broadcast scenario. For the example of Figure 3, the transport stream 300 may include the MPEG-2 transport stream which is a format internationally standardized in International Organization for Standardization (ISO)/MPEG.

[0043] Figure 4 illustrates an example layout of an MPEG2 transport stream for Internet Protocol television. Further, Figure 5 illustrates an example MPEG2 transport stream for Internet Protocol television.

[0044] With respect to Figures 4 and 5, in the streaming or progressive

download case over the Internet or cable, the delivery of audio/video may be implemented via a container file format (e.g., MPEG-4, avi, mkv, mov etc.). These containers may be incorporated in a transport stream for delivery over IPTV as depicted in Figures 4 and 5.

[0045] Figure 6 illustrates an example of MPEG2 transport stream multiplexing video, audio, and program information via a Program Map table.

[0046] Referring to Figure 6, an example of an MPEG2 transport stream 600 over Internet Protocol (IP) bitstream is illustrated, and depicts the Program Map table 602 in the stream.

[0047] Figure 7 illustrates an example MPEG2 transport stream Program Map table.

[0048] Referring to Figure 7, Figure 7 illustrates the Program Map table 602 from which content may be detected as either audio (i.e., non-a/v) or audio/video (i.e., a/v).

[0049] The transport stream may include four program specific informational tables that include Program Association Table (PAT), the Program Map table 602, Conditional Access Table (CAT), and Network Information Table (NIT). The Program Map table 602 may include information with respect to the program present in the transport stream, including the program number, and a list of the elementary streams that comprise the described MPEG-2 program. The Program Map table 602 may include locations for descriptors that describe the entire MPEG-2 program, as well as a descriptor for each elementary stream. Each elementary stream may be labeled with a stream type value. Figure 7 shows various stream type values stored in the Program Map table 602. If the stream involves audio (e.g., music, but no video), this may be detected via the audio frames that may be included in Audio Data Transport Stream (ADTS) (stream type: 15/0xF), described, for example, by ISO International Electrotechnical Commission (IEC) 13818-7.

[0050] Figure 8 illustrates an example of MPEG2 transport stream Advanced

Audio Coding-Audio Data Transport Stream.

[0051] Referring to Figure 8, a snapshot of ADTS from MPEG standard is shown in Figure 8 where the audio sync information is ascertained from the audio data and not from an audio/video sync timestamp (as in video-based audio) content. Advanced Audio Coding (AAC) based coding may represent a standard for data compression of music signals. Accordingly, ADTS may be used to discriminate between audio (e.g., music, but no video) and audio-for-video (non-audio). Additionally, stream type (2/0x2H) may be used to validate the video being present in the program (corresponding to music-video, TV show, or cinematic content for example). Thus the Program Map table 602 may be used to discriminate between audio (e.g., music, but no video) in ADTS or audio-for-video (e.g., moving images).

[0052] Figure 9 illustrates an example MPEG2 transport stream Program Map table with audio stream types identifying cinematic content.

[0053] Referring to Figure 9, identifiers for cinematic content (e.g., content either streamed or delivered through an external player) are shown in Figure 9 through audio element stream types (e.g., 128-194).

[0054] With respect to an audio-for-video program, an audio-for-video program may not be a movie (e.g., the audio-for-video program may be a television show in stereo or a music-video in stereo). Accordingly, heuristics may be applied under such conditions to extract additional program information from program metadata (e.g., duration of the audio-for-video program). The file-size and data rate may also be used to derive the duration of the program from the video or audio coding approach (e.g., H.264, H.265, AC-3, Advanced Audio Coding (AAC), etc.) depending on whether constant bitrate or variable bitrate coding is used. For example, for constant bitrate and variable bitrate coding, the duration (e.g., *d* in seconds) may be determined from audio coding as follows:

$$d_{CBR} = \frac{filesize \times 8}{bitrate} \qquad \text{Equation (1)}$$

$$d_{VBR} = \frac{N \times F}{f_s} \qquad \text{Equation (2)}$$

For Equations (1) and (2), *N* may represent the number of frames, *F* may represent samples/frame, f_s may represent the sampling frequency, *filesize* may be in kB (Kilobytes), and the bitrate may be in kbps (kilobits per sec). For cinematic clips, English movies may be 90 minutes or more, whereas television programs may not generally extend beyond 30 minutes, and music videos may be on an average approximately 4-5 minutes. Accordingly, downmixed cinematic content may be discriminated from television shows and music-videos. Additionally, a discriminant analysis (e.g., linear or pattern recognition techniques such as deep learning) may be applied to classify the content based on duration and the stream type data.

[0055] For example, Figure 10 illustrates an example content classifier 1000.

[0056] Referring to Figure 10, at block 1002, the content classifier 1000 may receive audio-video or audio from streaming or a terrestrial broadcast, or speech-keyword from a microphone. At block 1002, the content classifier 1000 may extract metadata, speech-keyword, or speech detection.

[0057] At block 1004, the content classifier 1000 may receive feature-vector of audio-video or audio metadata, or speech parameters (e.g., spectral centroid, fundamental frequency, formant 1 and 2, etc.). At block 1004, the content classifier 1000 may include a trained machine learning classifier (e.g., neural network, Bayesian classifier, Gaussian mixture model (GMM), clustering, etc.) to classify the content based on duration and the stream type data.

[0058] With respect to container formats (e.g., MP4 (MPEG-4), etc.), these formats may also be used for streaming over IP and hold both coded video and audio data. Since these formats may not be limited to storing audio data, these formats may be applicable for separating stereo (or downmixed cinematic audio) from multichannel cinematic audio using the techniques disclosed herein with respect to MP-2 transport stream. In this regard, the audio decoder parameters in the container may be analyzed to discriminate between multichannel audio and

stereo content.

[0059] Figure 11 illustrates an example of audio decoder element values indicating number of channels.

[0060] Referring to Figure 11, an example of AC-3 information is shown via the AC-3 standard associated with the Advanced Television Systems Committee (ATSC). With respect to Figure 11, the bitstream element at 1100 may identify the number of channels at 1102.

[0061] Figure 12 illustrates an example of an MPEG Advanced Audio Coding based bitstream syntax. Further, Figure 13 illustrates an example of MPEG Advanced Audio Coding based channel configurations.

[0062] Referring to Figure 12, an example of Advanced Audio Coding (AAC) based audio coding channel modes, also used for encoding cinematic content or sports/broadcast content, is shown in Figure 12. Specifically, Figure 12 illustrates the bitstream syntax 1200 where four bits (16 possible channel configurations) may be allocated to the channel configuration syntax at 1202 in the bitstream of audio specific configuration (ASC). In this regard, Figure 13 illustrates an example of MPEG Advanced Audio Coding based channel configurations. With respect to Figure 12, the containers may also include duration of the media embedded as metadata, and this metadata may reside in the header for streaming or progressive download. During audio encoding, additional parameters may be employed to determine the type of content. For example, cinematic and other audio with video content may use 48 kHz sampling frequency, whereas music may be sampled at 44.1 kHz. The bit-depth used for cinematic content may include 24 bits/sample representation, whereas broadcast content (including sports) may use 20 bits/sample, and whereas music may use 16 bits/sample representations.

[0063] According to an example, an MPEG-4 (MP4) may need to be packaged in a specific type of container, with the format for this container following the MPEG-4 Part 12 (ISO/IEC 14496-12) specification. Stream packaging may be

described as the process of making a multiplexed media file known as muxing, which combines multiple elements that enable control of the distribution delivery process into a single file. Some of these elements may be represented in self-contained atoms. An atom may be described as a basic data unit that contains a header and a data field. The header may include referencing metadata that describes how to find, process, and access the contents of the data field, which may include, for example, video frames, audio samples, interleaving AV data, captioning data, chapter index, title, poster, user data, and various technical metadata (e.g., coding scheme, timescale, version, preferred playback rate, preferred playback volume, movie duration, etc.).

[0064] In an MPEG-4 compliant container, every movie may include a {*moov*} atom. A movie atom may include a movie header atom (e.g., an mvhd atom) that defines the timescale and duration information for the entire movie, as well as its display characteristics. The movie atom may also contain a track atom (e.g., a trak atom) for each track in the movie. Each track atom may include one or more media atoms (e.g., an mdia atom) along with other atoms that define other track and movie characteristics. In this tree-like hierarchy, the moov atom may act as an index of the video data. The MPEG-4 muxer may store information about the file in the moov atom to enable the viewer to play and scrub the file as well. The file may not start to play until the player can access this index.

[0065] Unless specified otherwise, the moov atom may be stored at the end of the file in on-demand content, after all of the information describing the file has been generated. Depending on the type of on demand delivery technique selected (e.g., progressive download, streaming, or local playback), the location may be moved either to the end or to the beginning of the file.

[0066] If the planned delivery technique is progressive download or streaming (e.g., Real-Time Messaging Protocol (RTMP) or Hypertext Transfer Protocol (HTTP)), the moov atom may be moved to the beginning of the file. This ensures that the needed movie information may be downloaded first, enabling playback to

start. If the moov atom is located at the end of the file, the entire file may need to be downloaded before the beginning of playback. If the file is intended for local playback, then the location of the moov atom may not impact the start time, since the entire file is available for playback. The placement of the moov atom may be specified in various software packages through settings such as "progressive download," "fast start," "use streaming mode," or similar options. In this regard,

[0067] Using the duration information, heuristics may be used to parse whether the container includes music video or cinematic content. Further, the number of channels used may be determined from the decoder audio (e.g., channel configuration parameters).

[0068] According to an example, the MP-4 file-format may be used to discriminate between cinematic content and long-duration music content (e.g., live-concert audio-video recordings) using Object Content Information (OCI) which provides meta-information about objects. Object Content Information may define a set of descriptors and a stream type that have been defined in MPEG-4 Systems to carry information about the media object in general: Object Content Information descriptors and Object Content Information streams. Accordingly, a *ContentClassificationDescriptor* tag may be used by the creator or distributor, prior to encoding, for classification of the genre of the content. In this regard, Figure 14 illustrates an example of MP-4 metadata, where the metadata field may be extracted.

[0069] Voice in cinematic multichannel content may be located in the center channel and may be manipulated accordingly in terms of its preset. For business communications, voice/speech may be mono and a decoder output may trigger the appropriate preset.

[0070] With respect to content duration, statistics of feature film length and music video statistics may be obtained or derived from analysis. For example, Figure 15 illustrates an example of average duration of movie content along with some specific movies. In this regard, Figure 15 shows statistics of film length.

From Figure 15, the minimum average duration of cinematic content may be determined to be approximately 120 minutes.

[0071] Figure 16 illustrates an example distribution of a video length. Further, Figure 17 illustrates examples of distribution of various genres in seconds.

[0072] Referring to Figures 16 and 17, the statistics of music video files exhibit substantially smaller duration on an average.

[0073] The techniques disclosed herein to discriminate between multichannel cinematic (movie) content, stereo music, stereo downmixed movie, and other content based on transport streams, audio coding schemes, container formats, and parsing duration information may be used to apply specific genre tunings as disclosed herein.

[0074] With respect to preset selection and adaptation, based on the identification of the content genre, appropriate tuning presets may be applied to the content. According to examples, the presets may include the stereo music preset (e.g., for non-video and video-based music), voice preset (e.g., for entertainment), and movie preset (e.g., for stereo downmix). Additionally a fourth preset may be applied for multichannel, or next-generation audio (e.g., object-based audio or higher-order ambisonic based audio) for cinematic and entertainment content.

[0075] With respect to design and integration in personal devices, the techniques disclosed herein may be integrated, for example, in the any type of processors.

[0076] Figures 18-20 respectively illustrate an example block diagram 1800, an example flowchart of a method 1900, and a further example block diagram 2000 for audio tuning presets selection. The block diagram 1800, the method 1900, and the block diagram 2000 may be implemented on the apparatus 100 described above with reference to Figure 1 by way of example and not limitation. The block diagram 1800, the method 1900, and the block diagram 2000 may be practiced in other apparatus. In addition to showing the block diagram 1800, Figure 18 shows

hardware of the apparatus 100 that may execute the instructions of the block diagram 1800. The hardware may include a processor 1802, and a memory 1804 (i.e., a non-transitory computer readable medium) storing machine readable instructions that when executed by the processor cause the processor to perform the instructions of the block diagram 1800. The memory 1804 may represent a non-transitory computer readable medium. Figure 19 may represent a method for audio tuning presets selection, and the steps of the method. Figure 20 may represent a non-transitory computer readable medium 2002 having stored thereon machine readable instructions to provide audio tuning presets selection. The machine readable instructions, when executed, cause a processor 2004 to perform the instructions of the block diagram 2000 also shown in Figure 20.

[0077] The processor 1802 of Figure 18 and/or the processor 2004 of Figure 20 may include a single or multiple processors or other hardware processing circuit, to execute the methods, functions and other processes described herein. These methods, functions and other processes may be embodied as machine readable instructions stored on a computer readable medium, which may be non-transitory (e.g., the non-transitory computer readable medium 2002 of Figure 20), such as hardware storage devices (e.g., RAM (random access memory), ROM (read only memory), EPROM (erasable, programmable ROM), EEPROM (electrically erasable, programmable ROM), hard drives, and flash memory). The memory 1804 may include a RAM, where the machine readable instructions and data for a processor may reside during runtime.

[0078] Referring to Figures 1-18, and particularly to the block diagram 1800 shown in Figure 18, the memory 1804 may include instructions 1806 to determine whether content 104 included in a transport stream 106 includes stereo content 108 or multichannel content 110 by analyzing components 112 of the content 104 included in the transport stream 106, and a duration 114 of the content 104 included in the transport stream 106. For example, the analysis may include an analysis of a duration of the content extracted from the presentation timestamp

(PTS) metadata field at the beginning and at the end of the content, with the presentation timestamp being included in the transport stream.

[0079] The processor 1802 may fetch, decode, and execute the instructions 1808 to, in response to a determination that the content 104 included in the transport stream 106 includes the stereo content 108, apply a stereo content preset 120 to the content 104 included in the transport stream 106.

[0080] The processor 1802 may fetch, decode, and execute the instructions 1810 to, in response to a determination that the content 104 included in the transport stream 106 includes the multichannel content 110, apply a multichannel content preset 122 to the content 104 included in the transport stream 106.

[0081] Referring to Figures 1-17 and 19, and particularly Figure 19, for the method 1900, at block 1902, the method may include determining whether content 104 included in a transport stream 106 includes stereo content 108 or multichannel content 110 by analyzing components 112 of the content 104 included in the transport stream 106, and a duration 114 of the content 104 included in the transport stream 106 by analyzing a file-size and a data rate associated with the content 104 included in the transport stream 106. Other techniques including determining the duration from the timestamp may be used.

[0082] At block 1904, in response to a determination that the content 104 included in the transport stream 106 includes the stereo content 108, the method may include applying a stereo content preset 120 to the content 104 included in the transport stream 106.

[0083] At block 1906, in response to a determination that the content 104 included in the transport stream 106 includes the multichannel content 110, the method may include applying a multichannel content preset 122 to the content 104 included in the transport stream 106.

[0084] Referring to Figures 1-17 and 20, and particularly Figure 20, for the block diagram 2000, the non-transitory computer readable medium 2002 may

include instructions 2006 to determine whether content 104 included in a transport stream 106 includes stereo content 108 or multichannel content 110 by analyzing components 112 of the content 104 included in the transport stream 106, and a duration 114 of the content 104 included in the transport stream 106 by comparing the duration 114 to predetermined durations for different types of stereo content 108 and multichannel content 110.

[0085] The processor 2004 may fetch, decode, and execute the instructions 2008 to, in response to a determination that the content 104 included in the transport stream 106 includes the stereo content 108, apply a stereo content preset 120 to the content 104 included in the transport stream 106.

[0086] The processor 2004 may fetch, decode, and execute the instructions 2010 to, in response to a determination that the content 104 included in the transport stream 106 includes the multichannel content 110, apply a multichannel content preset 122 to the content 104 included in the transport stream 106.

[0087] What has been described and illustrated herein is an example along with some of its variations. The terms, descriptions and figures used herein are set forth by way of illustration only and are not meant as limitations. Many variations are possible within the spirit and scope of the subject matter, which is intended to be defined by the following claims -- and their equivalents -- in which all terms are meant in their broadest reasonable sense unless otherwise indicated.

What is claimed is:

1. An apparatus comprising:

a processor; and

a non-transitory computer readable medium storing machine readable instructions that when executed by the processor cause the processor to:

determine whether content included in a transport stream includes stereo content or multichannel content by analyzing

components of the content included in the transport stream, and

a duration of the content included in the transport stream, wherein the components of the content included in the transport stream are included in a container included in the transport stream;

in response to a determination that the content included in the transport stream includes the stereo content, apply a stereo content preset to the content included in the transport stream; and

in response to a determination that the content included in the transport stream includes the multichannel content, apply a multichannel content preset to the content included in the transport stream.

2. The apparatus according to claim 1, wherein for the stereo content, the instructions are further to cause the processor to:

determine whether the stereo content includes a video, whether the stereo content does not include the video, and whether the stereo content is downmixed cinematic content; and

in response to a determination that the stereo content includes the video, that the stereo content does not include the video, or that the stereo content is

downmixed cinematic content, apply a corresponding type of the stereo content preset to the content included in the transport stream.

3. The apparatus according to claim 1, the instructions are further to cause the processor to:

determine whether the content included in the transport stream includes voice in Voice over Internet Protocol (VoIP), or speech; and

in response to a determination that the content included in the transport stream includes the voice in VoIP, or the speech, apply a voice preset to the content included in the transport stream.

4. The apparatus according to claim 1, wherein the instructions are further to cause the processor to:

determine whether the content included in the transport stream includes microphone-captured speech; and

in response to a determination that the content included in the transport stream includes the microphone-captured speech, apply a microphone-captured speech voice preset to the content included in the transport stream.

5. The apparatus according to claim 1, wherein the instructions are further to cause the processor to:

determine, based on the analysis of the components in a Program Map table included in the container, whether the content included in the transport stream includes audio content, or audio and video content, wherein the components include audio frames for the audio content; and

in response to a determination that the content included in the transport

stream includes the audio content, or the audio and video content, selectively apply the stereo content preset or the multichannel content preset to the content included in the transport stream.

6. The apparatus according to claim 1, wherein the instructions are further to cause the processor to:

determine, based on the analysis of the components in a Program Map table included in the container, whether the content included in the transport stream includes audio content, or audio-for-video content; and

in response to a determination that the content included in the transport stream includes the audio content, or the audio-for-video content, selectively apply the stereo content preset or the multichannel content preset to the content included in the transport stream.

7. The apparatus according to claim 1, wherein the instructions are further to cause the processor to:

determine the duration of the content included in the transport stream by analyzing a file-size and a data rate associated with the content included in the transport stream.

8. The apparatus according to claim 7, wherein the data rate includes a constant bitrate or a variable bitrate.

9. The apparatus according to claim 1, wherein the instructions are further to cause the processor to:

analyze the duration of the content included in the transport stream by

comparing the duration to predetermined durations for different types of stereo content and multichannel content.

10. A method comprising:

determining, by a processor, whether content included in a transport stream includes stereo content or multichannel content by analyzing

components of the content included in the transport stream, and

a duration of the content included in the transport stream, wherein

the duration is analyzed by analyzing a file-size and a data rate associated with the content included in the transport stream, and

the components of the content included in the transport stream are included in a container included in the transport stream;

in response to a determination that the content included in the transport stream includes the stereo content, applying a stereo content preset to the content included in the transport stream; and

in response to a determination that the content included in the transport stream includes the multichannel content, applying a multichannel content preset to the content included in the transport stream.

11. The method according to claim 10, wherein the components of the content included in the transport stream include an audio coding scheme that specifies a type of the content included in the transport stream.

12. The method according to claim 10, wherein the components of the content included in the transport stream include a container format that is used to determine whether the content included in the transport stream includes the stereo

content or the multichannel content.

13. A non-transitory computer readable medium having stored thereon machine readable instructions, the machine readable instructions, when executed, cause a processor to:

determine whether content included in a transport stream includes stereo content or multichannel content by analyzing

components of the content included in the transport stream, and

a duration of the content included in the transport stream, wherein

the duration is analyzed by comparing the duration to predetermined durations for different types of stereo content and multichannel content;

in response to a determination that the content included in the transport stream includes the stereo content, apply a stereo content preset to the content included in the transport stream; and

in response to a determination that the content included in the transport stream includes the multichannel content, apply a multichannel content preset to the content included in the transport stream.

14. The non-transitory computer readable medium according to claim 13, wherein the instructions are further to cause the processor to:

determine, based on the analysis of the components in a Program Map table included in a container included in the transport stream, whether the content included in the transport stream includes audio content, or audio and video content, wherein the components include audio frames for the audio content; and

in response to a determination that the content included in the transport

stream includes the audio content, or the audio and video content, selectively apply the stereo content preset or the multichannel content preset to the content included in the transport stream.

15. The non-transitory computer readable medium according to claim 13, wherein the instructions are further to cause the processor to:

determine, based on the analysis of the components in a Program Map table included in a container included in the transport stream, whether the content included in the transport stream includes audio content, or audio-for-video content; and

in response to a determination that the content included in the transport stream includes the audio content, or the audio-for-video content, selectively apply the stereo content preset or the multichannel content preset to the content included in the transport stream.

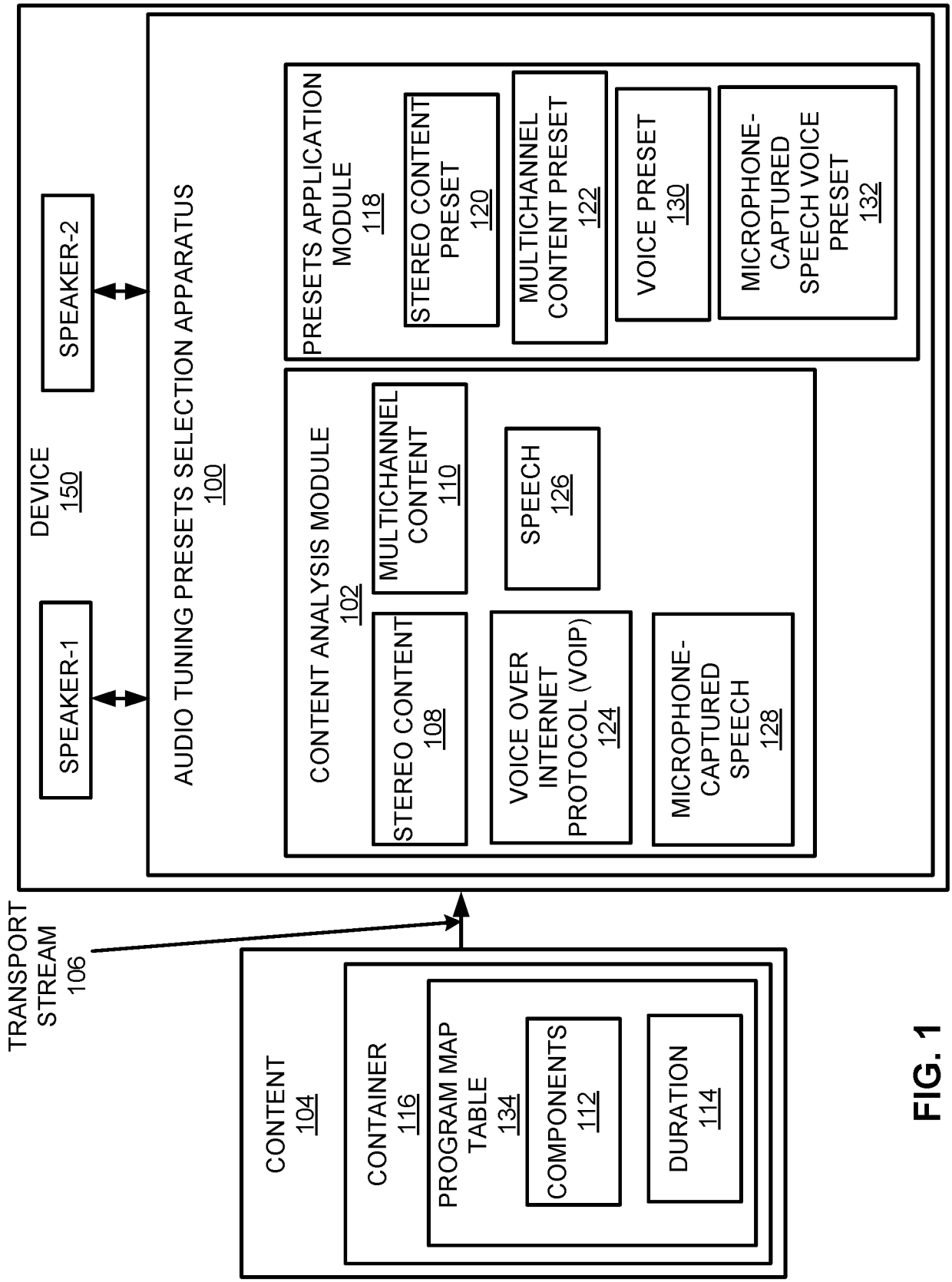

TRANSPORT STREAM 106
CONTENT 104
CONTAINER 116
PROGRAM MAP TABLE 134
COMPONENTS 112
DURATION 114
DEVICE 150
SPEAKER-1
SPEAKER-2
AUDIO TUNING PRESETS SELECTION APPARATUS 100
CONTENT ANALYSIS MODULE 102
STEREO CONTENT 108
MULTICHANNEL CONTENT 110
VOICE OVER INTERNET PROTOCOL (VOIP) 124
SPEECH 126
MICROPHONE-CAPTURED SPEECH 128
PRESETS APPLICATION MODULE 118
STEREO CONTENT PRESET 120
MULTICHANNEL CONTENT PRESET 122
VOICE PRESET 130
MICROPHONE-CAPTURED SPEECH VOICE PRESET 132

FIG. 1

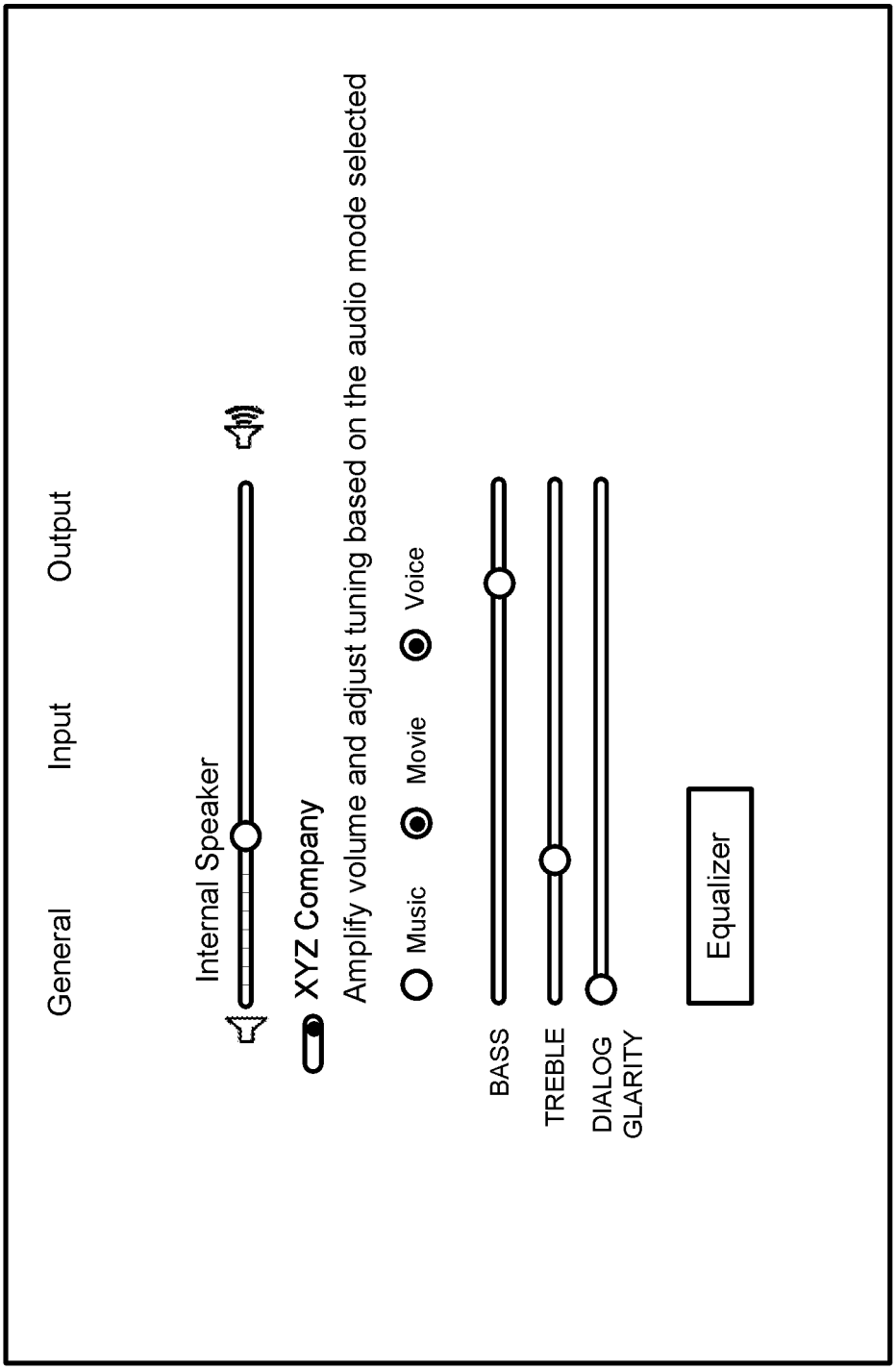

General
Input
Output
Internal Speaker
XYZ Company
Amplify volume and adjust tuning based on the audio mode selected
Music
Movie
Voice
BASS
TREBLE
DIALOG GLARITY
Equalizer

FIG. 2

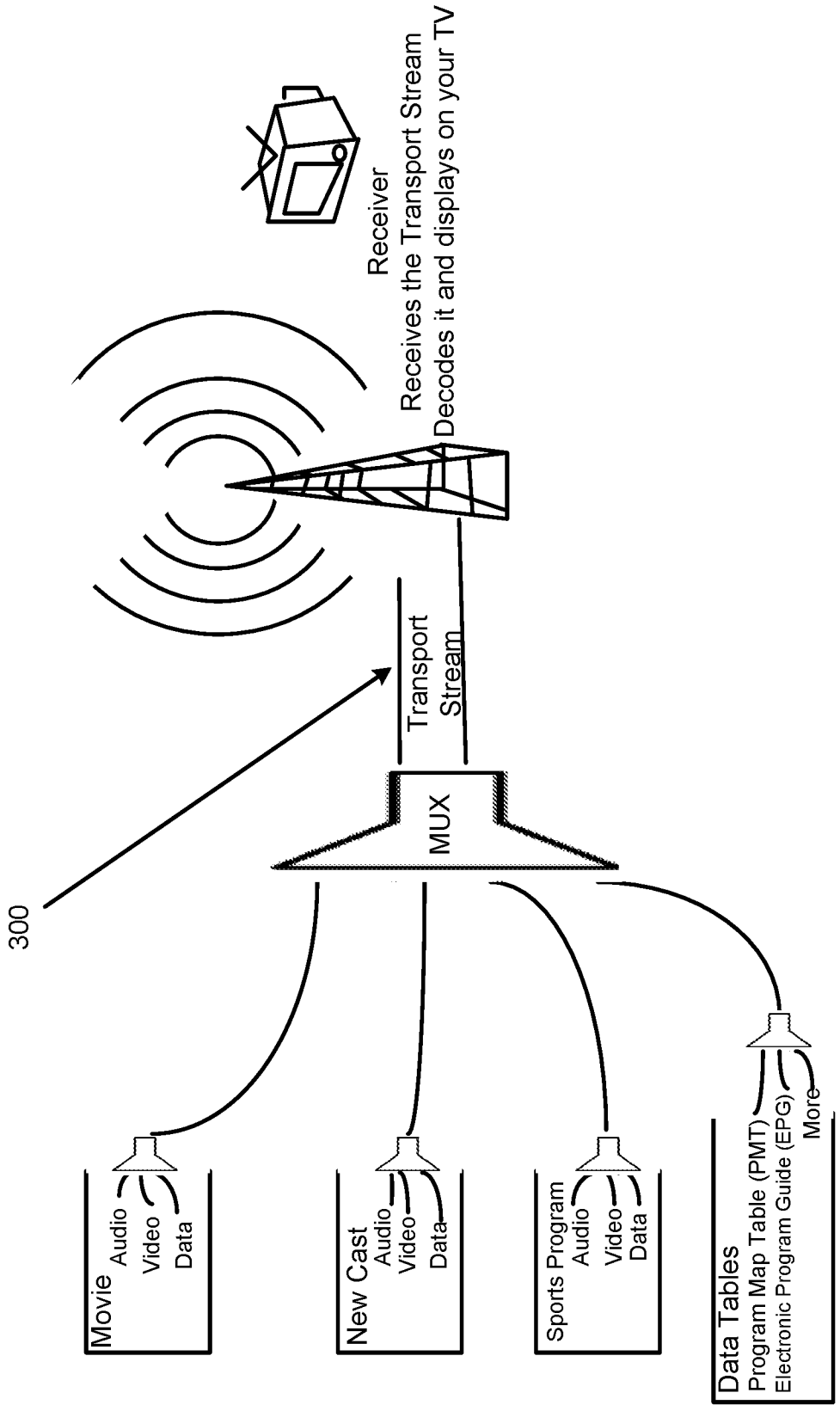

300
MUX
Transport Stream
Receiver
Receives the Transport Stream
Decodes it and displays on your TV
Movie
Audio
Video
Data
New Cast
Audio
Video
Data
Sports Program
Audio
Video
Data
Data Tables
Program Map Table (PMT)
Electronic Program Guide (EPG)
More

FIG. 3

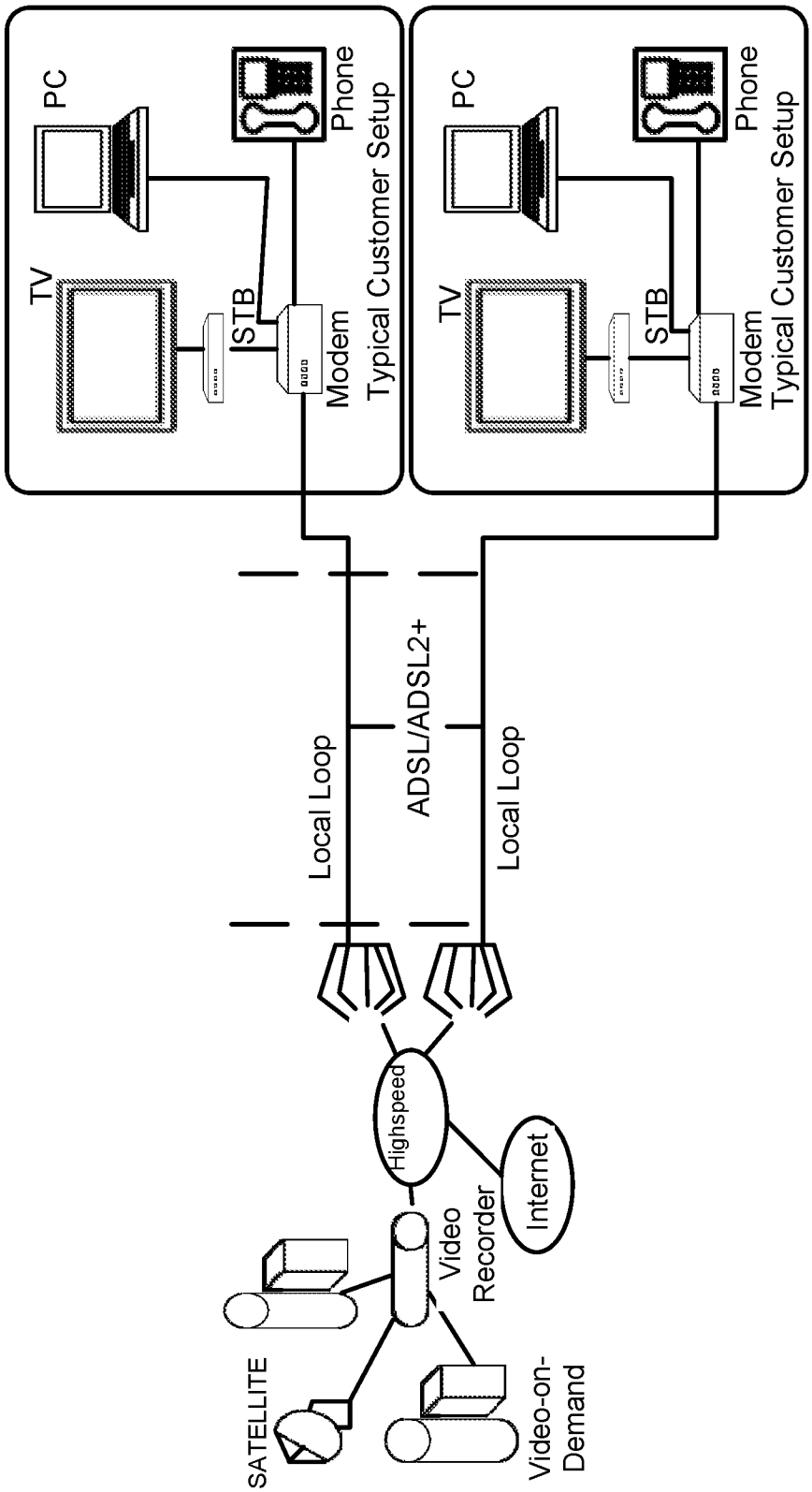

PC
Phone
TV
STB
Modem
Typical Customer Setup
PC
Phone
TV
STB
Modem
Typical Customer Setup
Local Loop
ADSL/ADSL2+
Local Loop
Highspeed
Internet
Video
Recorder
SATELLITE
Video-on-
Demand

FIG. 4

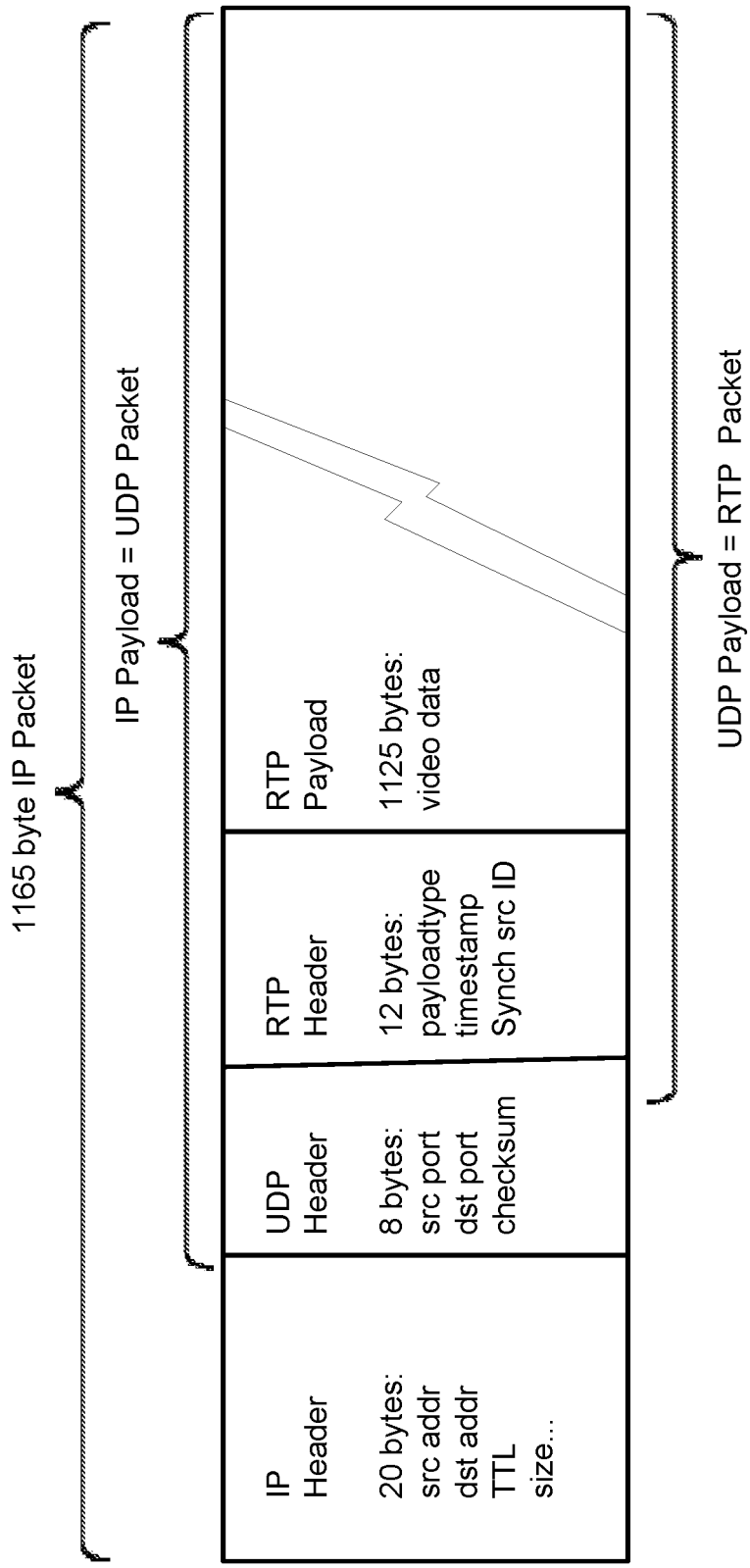

1165 byte IP Packet
IP Payload = UDP Packet
UDP Payload = RTP Packet
IP Header
20 bytes:
src addr
dst addr
TTL
size...
UDP Header
8 bytes:
src port
dst port
checksum
RTP Header
12 bytes:
payloadtype
timestamp
Synch src ID
RTP Payload
1125 bytes:
video data

FIG. 5

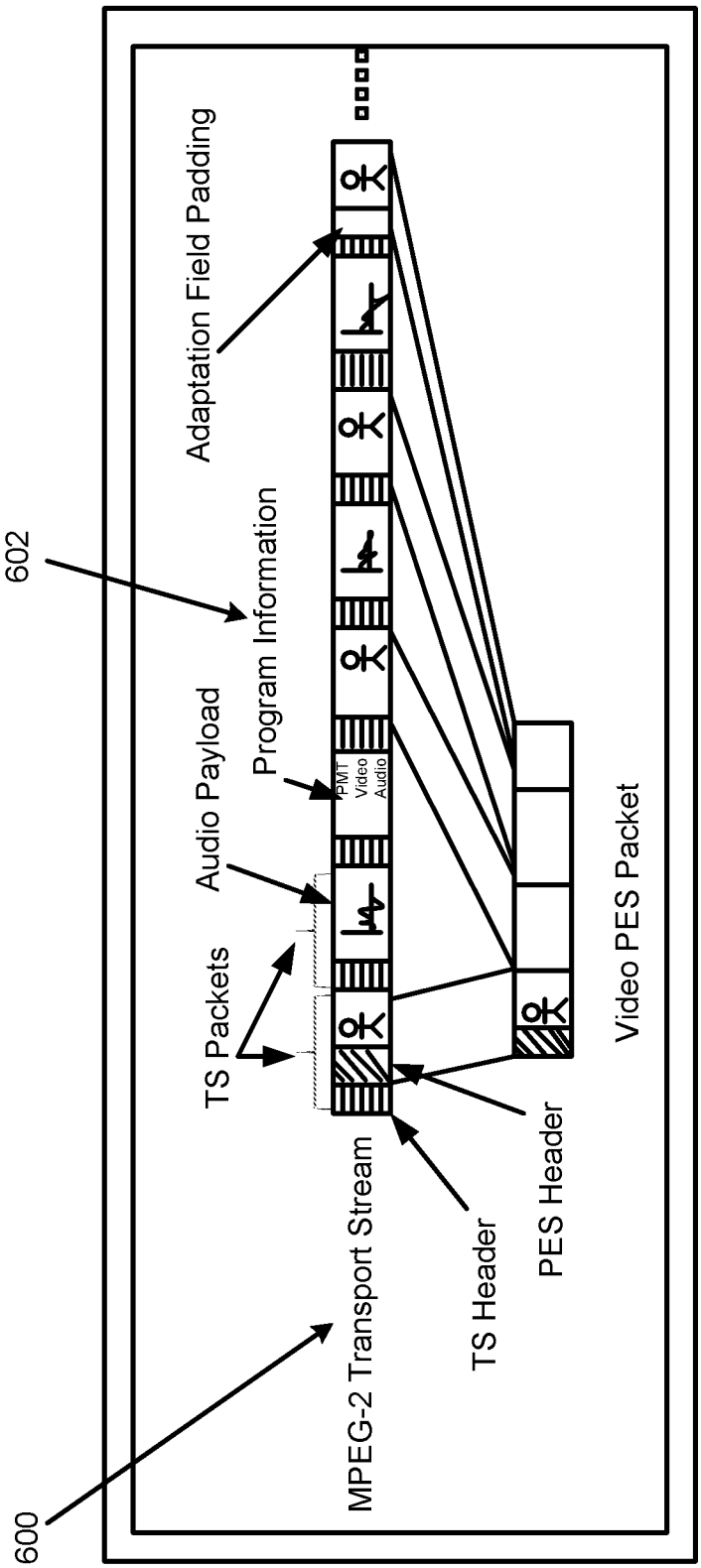

600
602
MPEG-2 Transport Stream
TS Packets
Audio Payload
Program Information
Adaptation Field Padding
PMT
Video
Audio
TS Header
PES Header
Video PES Packet

FIG. 6

usable MPEG-2 stream_type values

Stream_type (dec/hex)	Description
1/0x1	ISO/IEC 11172 Video (MPEG-1)
2/0x2	ITU-T Rec. H.262 I ISO/IEC 13818-2 (MPEG-2) Video or ISO/IEC 11172-2 (MPEG-1) constrained parameter video stream
3/0x3	ISO/IEC 11172 Audio (MPEG-1)
4/0x4	ISO/IEC 13818-3 Audio (MPEG-2)
5/0x5	MPEG-2 private table sections
6/0x6	MPEG-2 Packetized Elementary Stream packets containing private data
7/0x7	MHEG Packets
8/0x8	MPEG-2 Annex A DSM CC
9/0x9	ITU-T Rec. H.222.1
10/0xA	ISO/IEC 13818-6 Type A
11/0xB	ISO/IEC 13818-6 Type B
12/0xC	ISO/IEC 13818-6 Type C
13/0xD	ISO/IEC 13818-6 Type D
14/0xE	ISO/IEC 13818-1 (MPEG-2) auxiliary
15/0xF	ISO/IEC 13818-7 Audio with ADTS transport syntax
16/0x10	ISO/IEC 14496-2 (MPEG-4) Visual
17/0x11	ISO/IEC 14496-3 Audio with the LATM transport syntax as defined in ISO/IEC 14496-3/AMD 1
18/0x12	ISO/IEC 14496-1 SL-packetized stream or FlexMux stream carried in PES packets
19/0x13	ISO/IEC 14496-1 SL-packetized stream or FlexMux stream carried in ISO/IEC14496_sections.
20/0x14	ISO/IEC 13818-6 Synchronized Download Protocol
21/0x15	Metadata carried in PES packets
22/0x16	Metadata carried in metadata_sections
128-255 0x80-0xFF	MPEG-2 User Private
129/0x81	ATSC AC-3 Audio

602

FIG. 7

Protocol/Format	Type	Defined by	Features	Typical Use
ADTS (Audio Data Transport Stream)	Transport Audio Only	MPEG-2 ISO/IEC 13818-7[1] (also in MPEG-4)	Self-contained bitstream with Sync information, originating From MPEG-2 AAC.	Transport for MPEG-2 AAC Sometimes used for mobile handsets and devices, since parsing MP4FF has been quite demanding memory-wise in the past. Some mobile phones that are able to play AAC-LC or HE-AAC support only ADTS.

FIG. 8

22	0x16	Sectioned metada
23	0x17	ISO/IEC 13818-6 DSM CC Data Carousel metadata
24	0x18	ISO/IEC 13818-6 DSM CC Object Carousel metadata
25	0x19	ISO/IEC 13818-6 Synchronized Download Protocol metadata
26	0x1A	ISO/IEC 13818-11 IPMP
27	0x1B	ITU-T Rec.H.264 and ISO/IEC 14496-10 (lower bit-rate video) in a packetized stream
28-35	0x1C-0x23	Reserved.
36	0x24	ITU-T Rec.H.265 and ISO/IEC 23008-2 (Ultra HD video) in a packetized stream
37-65	0x25-0x41	Reserved.
66	0x42	Chinese Video Standard in a packetized stream
67-127	0x43-0x7F	Reserved.
128	0x80	ITU-T Rec.H.262 and ISO/IEC 13818-2 for DigiCipher II or PCM audio
129	0x81	channel audio
130	0x82	SCTE subtitle or DTS 6 channel audio
131	0x83	audio
132	0x84	channel audio
133	0x85	8 channel audio
134	0x86	SCTE-35 digital program insertion cue message DTS 8 channel lossless audio
135	0x87	16 channel audio
136-143	0x88-0x8F	Privately defined.
144	0x90	Presentation Graphic Stream (subtitling) in a packetized stream
145	0x95	ATSC DSM CC Network Resources table
146-191	0x95-0xDF	Privately defined.

FIG. 9

192	0xC0	DigiCipher II text in a packetized stream
193	0xC1	channel audio with AES-128-CBC data encryption In a packetized stream
194	0xC2	ATSC DSM CC synchronous data
195-206	0xC3-0xCE	Privately defined.
207	0xCF	ISO/IEC 13818-7 ADTS AAC with AES-128-CBC frame encryption in a packetized stream
208	0xD0	Privately defined.
209	0xD1	BBC Dirac (Ultra HD video) in a packetized stream
210-218	0xD2-0xDA	Privately defined.
219	0xDB	ITU-T Rec. H.264 and ISO/IEC 14496-10 with AES-128-CBC slice encryption In a packetized stream
220-233	0xDC-0xE9	Privately defined.
234	0xEA	Video 9 (lower bit-rate video) In a packetized stream
235-255	0xEB-0xFF	Privately defined.

FIG. 9
(CONT)

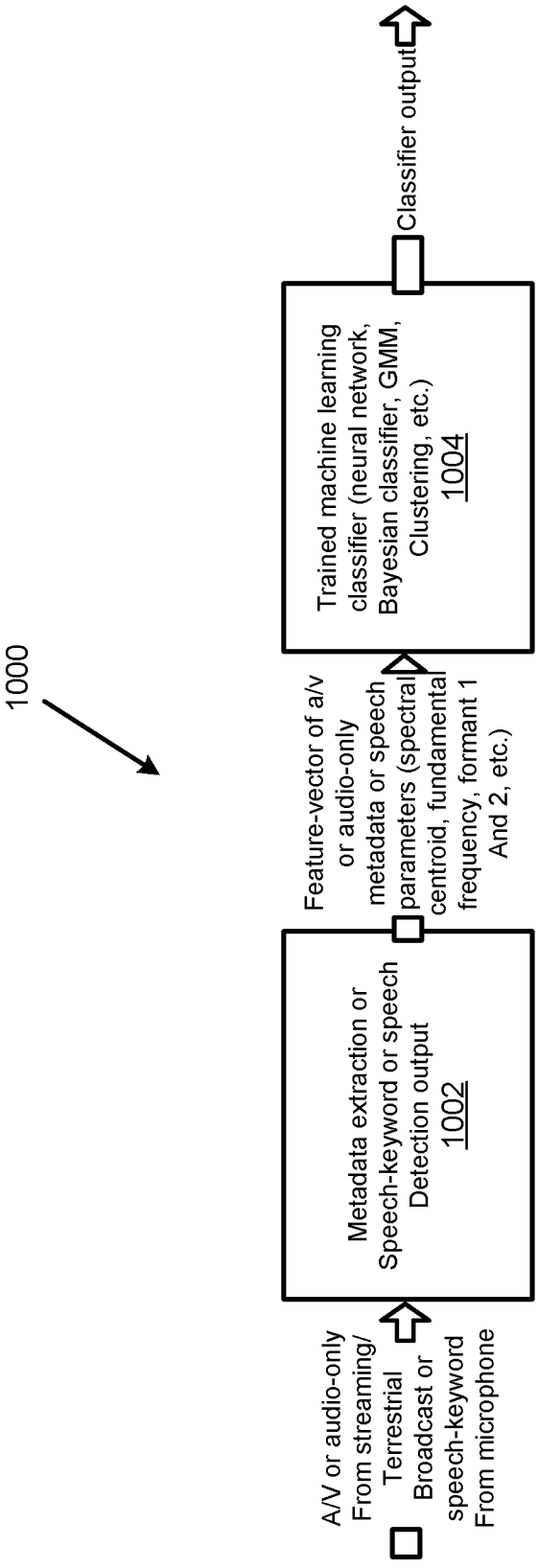

1000
A/V or audio-only
From streaming/
Terrestrial
Broadcast or
speech-keyword
From microphone
Metadata extraction or
Speech-keyword or speech
Detection output
1002
Feature-vector of a/v
or audio-only
metadata or speech
parameters (spectral
centroid, fundamental
frequency, formant 1
And 2, etc.)
Trained machine learning
classifier (neural network,
Bayesian classifier, GMM,
Clustering, etc.)
1004
Classifier output

FIG. 10

1100

1102

acmod	Audio Coding Mode	nfchans	Channel Array Ordering
'000'	1+1	2	Ch1, Ch2
'001'	1/0	1	C
'010'	2/0	2	L, R
'011'	3/0	3	L, C, R
'100'	2/1	3	L, R, S
'101'	3/1	4	L, C, R, S
'110'	2/2	4	L, R, SL, SR
'111'	3/2	5	L, C, R, SL, SR

FIG. 11

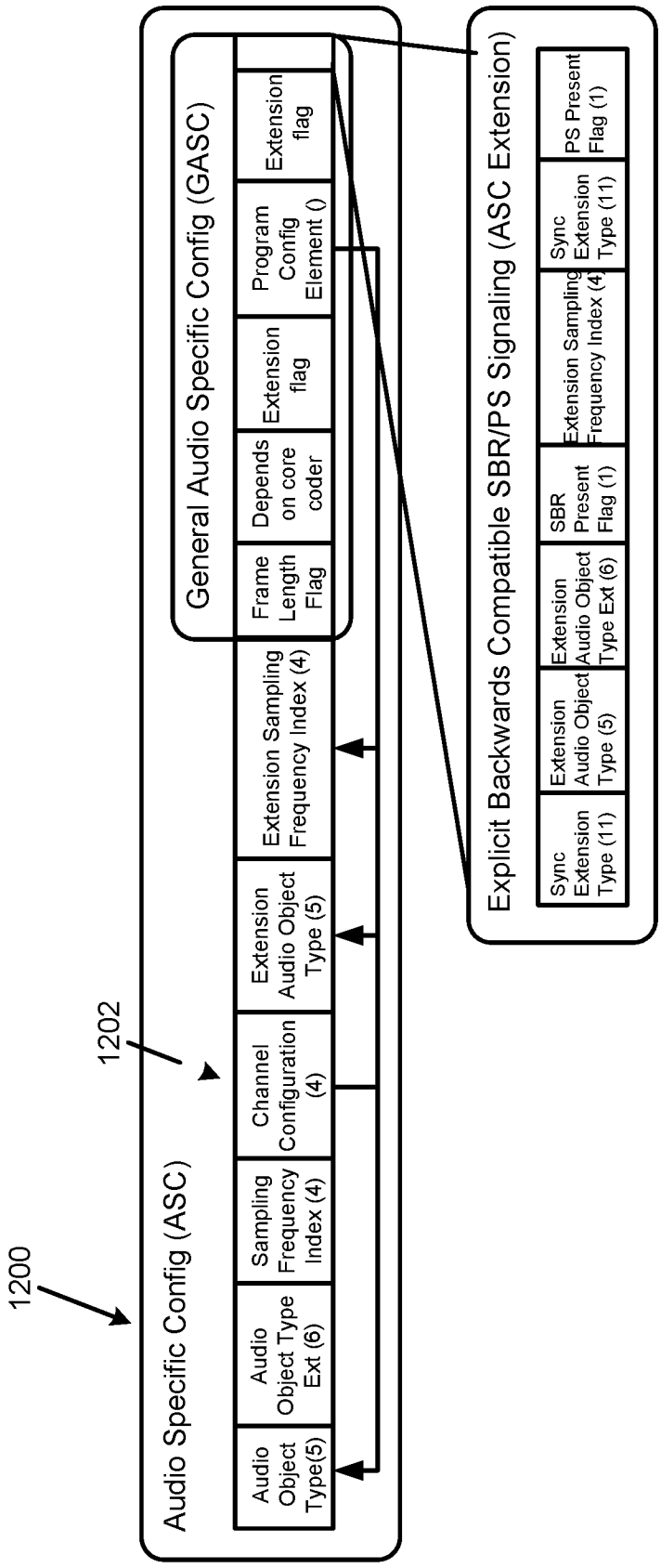

1200
Audio Specific Config (ASC)
Audio Object Type(5)
Audio Object Type Ext (6)
Sampling Frequency Index (4)
Channel Configuration (4)
Extension Audio Object Type (5)
Extension Sampling Frequency Index (4)
1202
General Audio Specific Config (GASC)
Frame Length Flag
Depends on core coder
Extension flag
Program Config Element ()
Extension flag
Explicit Backwards Compatible SBR/PS Signaling (ASC Extension)
Sync Extension Type (11)
Extension Audio Object Type (5)
Extension Audio Object Type Ext (6)
SBR Present Flag (1)
Extension Sampling Frequency Index (4)
Sync Extension Type (11)
PS Present Flag (1)

FIG. 12

Channel configuration	included channel element	channel to speaker mapping
0	-	
1	single_channel_element()	center front speaker
2	channel_pair_element()	left, right front speakers
3	single_channel_element() channel_pair_element()	center front speaker left, right front speakers
4	single_channel_element() channel_pair_element() single_channel_element()	center front speaker left, right center front speakers rear surround speakers
5	single_channel_element() channel_pair_element() channel_pair_element()	center front speaker left, right front speakers left surround, right surround speakers
6	single_channel_element() channel_pair_element() channel_pair_element() 1fe_channel_element()	center front speaker left, right front speakers left surround, right surround speakers, low frequency effects speaker
7	single_channel_element() channel_pair_element() channel_pair_element() channel_pair_element() 1fe_channel_element()	center front speaker left, right front center speakers, left, right front speakers left surround, right surround speakers, low frequency effects speaker
8-10	-	reserved

FIG. 13

Channel configuration	included channel element	channel to speaker mapping
11	single_channel_element(), channel_pair_element(), channel_pair_element(), single_channel_element(), 1fe_element()	center front speaker left, right front speakers, left surround, right surround speakers, rear center speaker, low frequency enhancement speaker
12	single_channel_element(), channel_pair_element(), channel_pair_element(), channel_pair_element(), 1fe_element()	center front speaker left, right front speakers, left surround, right surround speakers, rear surround left, right speakers, low frequency enhancement speaker
14	single_channel_element(), channel_pair_element(), channel_pair_element(), 1fe_element(), channel_pair_element()	center front speaker left, right front speakers, left surround, right surround speakers, low frequency enhancement speaker, Left, right front vertical height speakers
15	-	reserved

FIG. 13
(CONT)

Key	Field	Low-level identifier
“title”	Name	‘\251nam’
“author”	Artist	‘\251ART’
“album_artist”	Album Artist	‘aART’
“album”	Album	‘\251alb’
“grouping”	Grouping	‘\251grp’
“composer”	Composer	‘\251wrt’
“year”	Year	‘\251day’
“track”	Track Number	‘trkn’
“comment”	Comments	‘\251cmt’
“genre”	Genre	‘\251gen’
“copyright”	??	‘\251cpy’
“description”	Description	‘desc’
“synopsis”	Information dialog when selecting “Show Description” in context menu	‘ldes’
“show”	Show	‘tvsh’
“episode_id”	Episode ID	‘tven’
“network”	??	‘tvnn’
“lyrics”	Lyrics	‘\251lyr’

FIG. 14

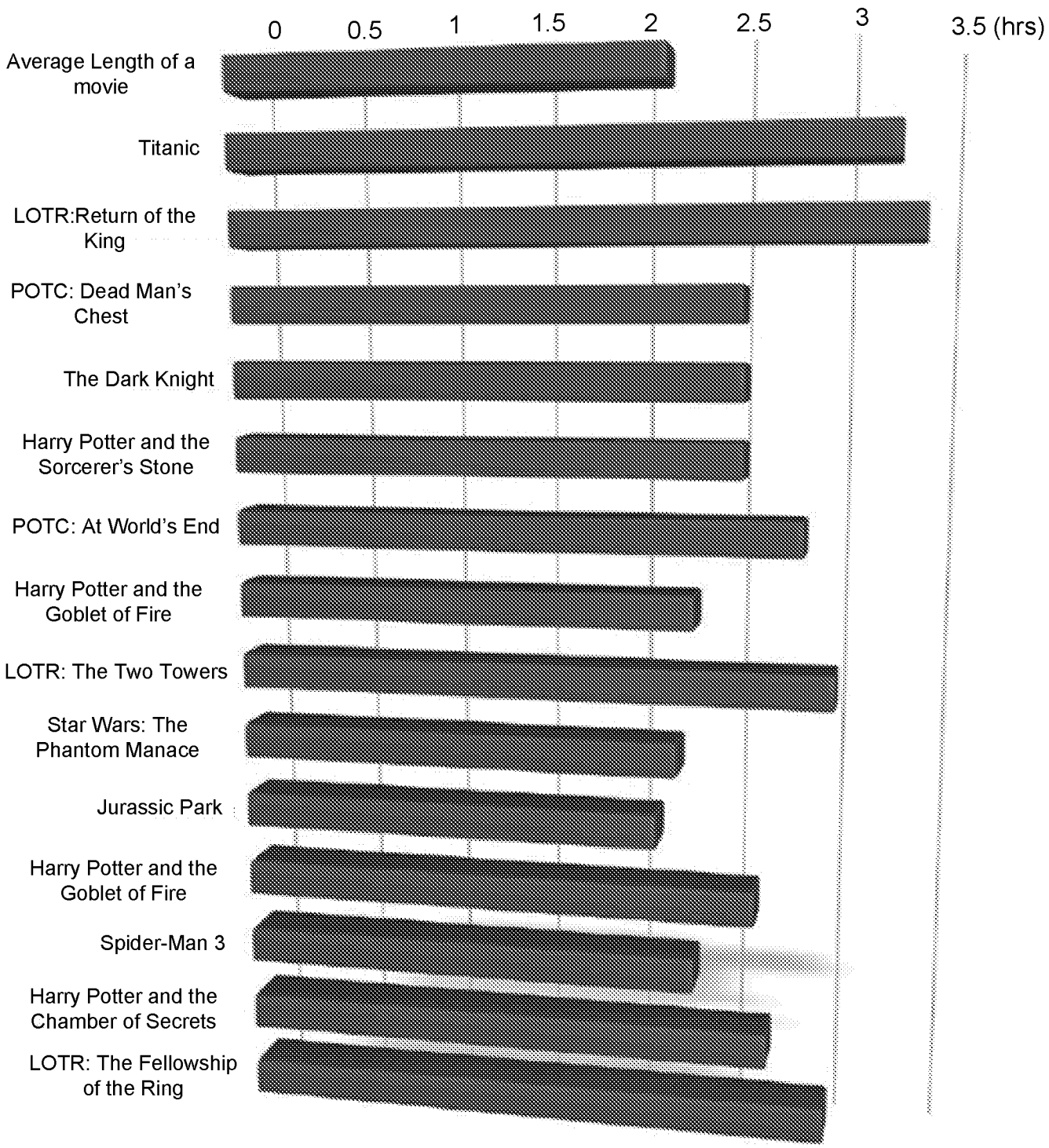

0
0.5
1
1.5
2
2.5
3
3.5 (hrs)
Average Length of a movie
Titanic
LOTR:Return of the King
POTC: Dead Man's Chest
The Dark Knight
Harry Potter and the Sorcerer's Stone
POTC: At World's End
Harry Potter and the Goblet of Fire
LOTR: The Two Towers
Star Wars: The Phantom Manace
Jurassic Park
Harry Potter and the Goblet of Fire
Spider-Man 3
Harry Potter and the Chamber of Secrets
LOTR: The Fellowship of the Ring

FIG. 15

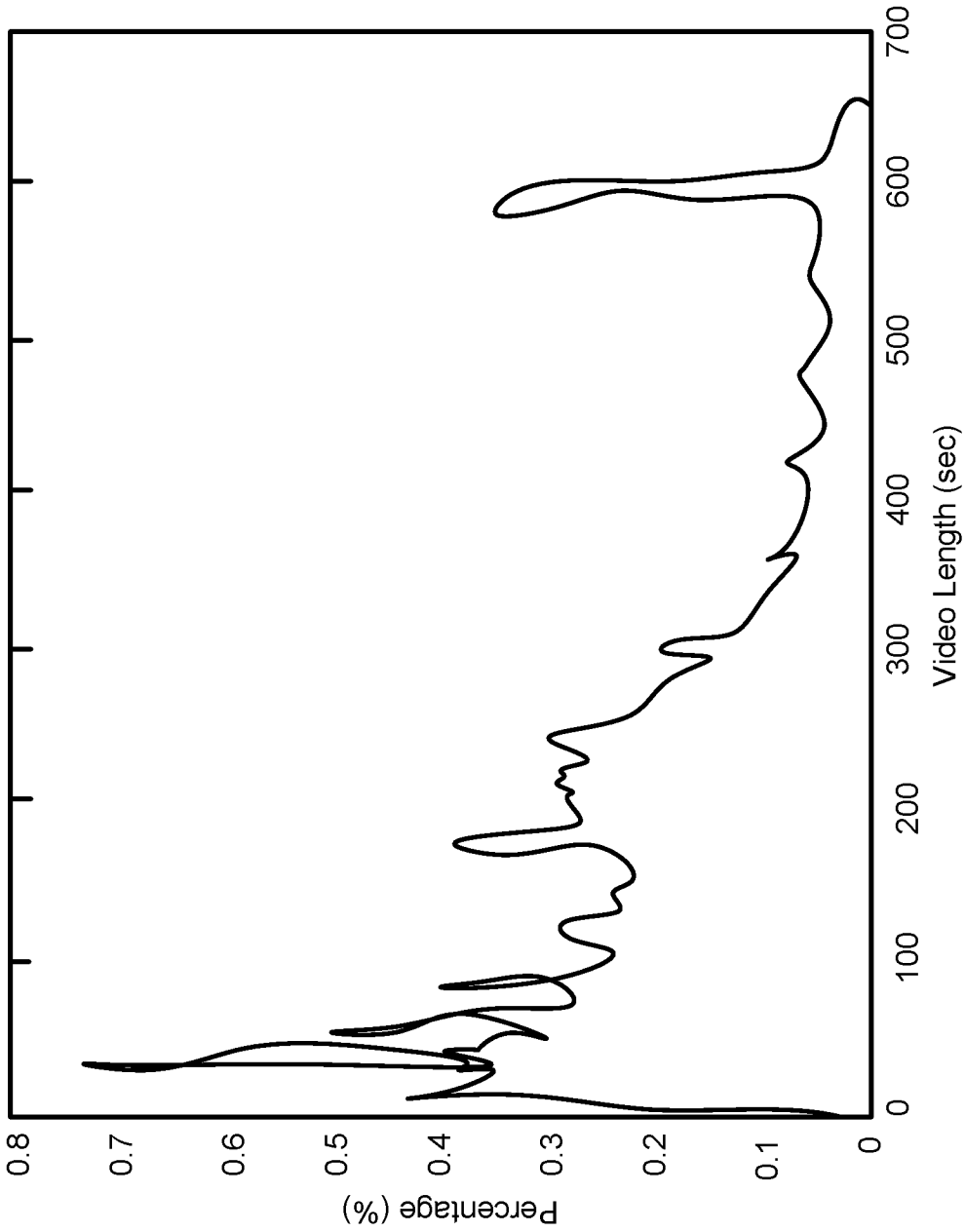

Percentage (%)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
Video Length (sec)
0
100
200
300
400
500
600
700

FIG. 16

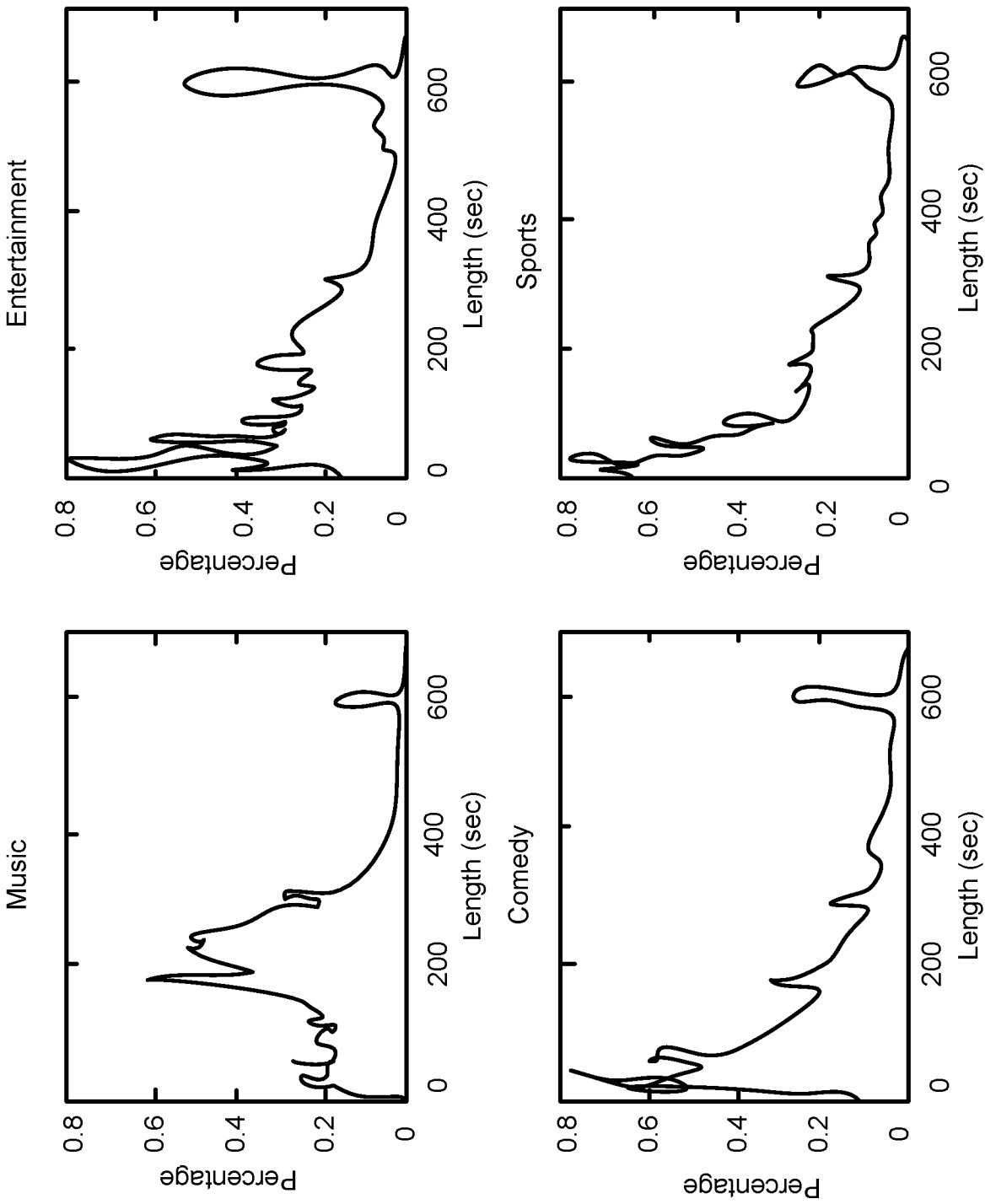

Entertainment
Sports
Music
Comedy
Percentage
Length (sec)
0
0.2
0.4
0.6
0.8
0
200
400
600

FIG. 17

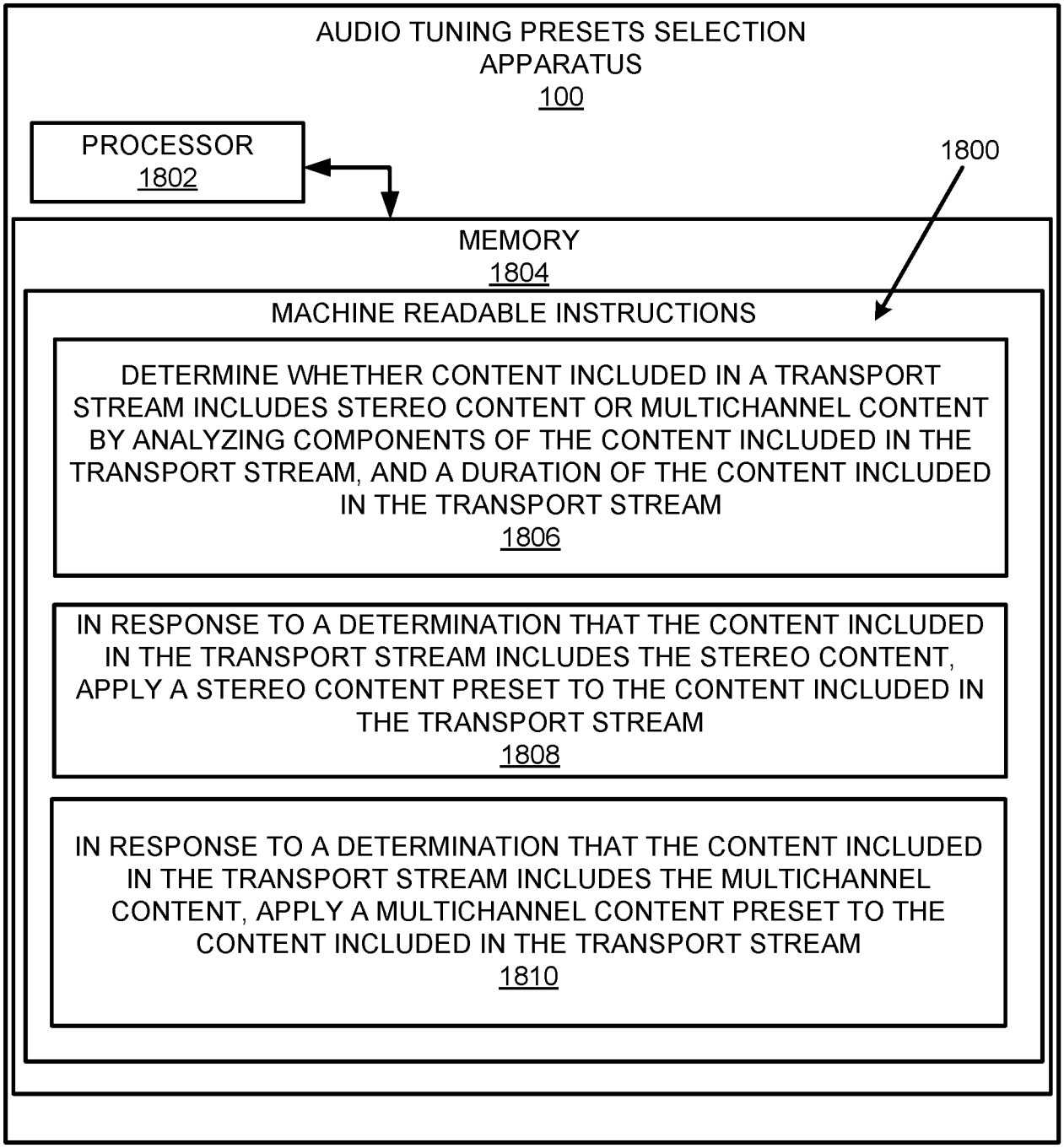

AUDIO TUNING PRESETS SELECTION APPARATUS
100
PROCESSOR
1802
1800
MEMORY
1804
MACHINE READABLE INSTRUCTIONS
DETERMINE WHETHER CONTENT INCLUDED IN A TRANSPORT STREAM INCLUDES STEREO CONTENT OR MULTICHANNEL CONTENT BY ANALYZING COMPONENTS OF THE CONTENT INCLUDED IN THE TRANSPORT STREAM, AND A DURATION OF THE CONTENT INCLUDED IN THE TRANSPORT STREAM
1806
IN RESPONSE TO A DETERMINATION THAT THE CONTENT INCLUDED IN THE TRANSPORT STREAM INCLUDES THE STEREO CONTENT, APPLY A STEREO CONTENT PRESET TO THE CONTENT INCLUDED IN THE TRANSPORT STREAM
1808
IN RESPONSE TO A DETERMINATION THAT THE CONTENT INCLUDED IN THE TRANSPORT STREAM INCLUDES THE MULTICHANNEL CONTENT, APPLY A MULTICHANNEL CONTENT PRESET TO THE CONTENT INCLUDED IN THE TRANSPORT STREAM
1810

FIG. 18

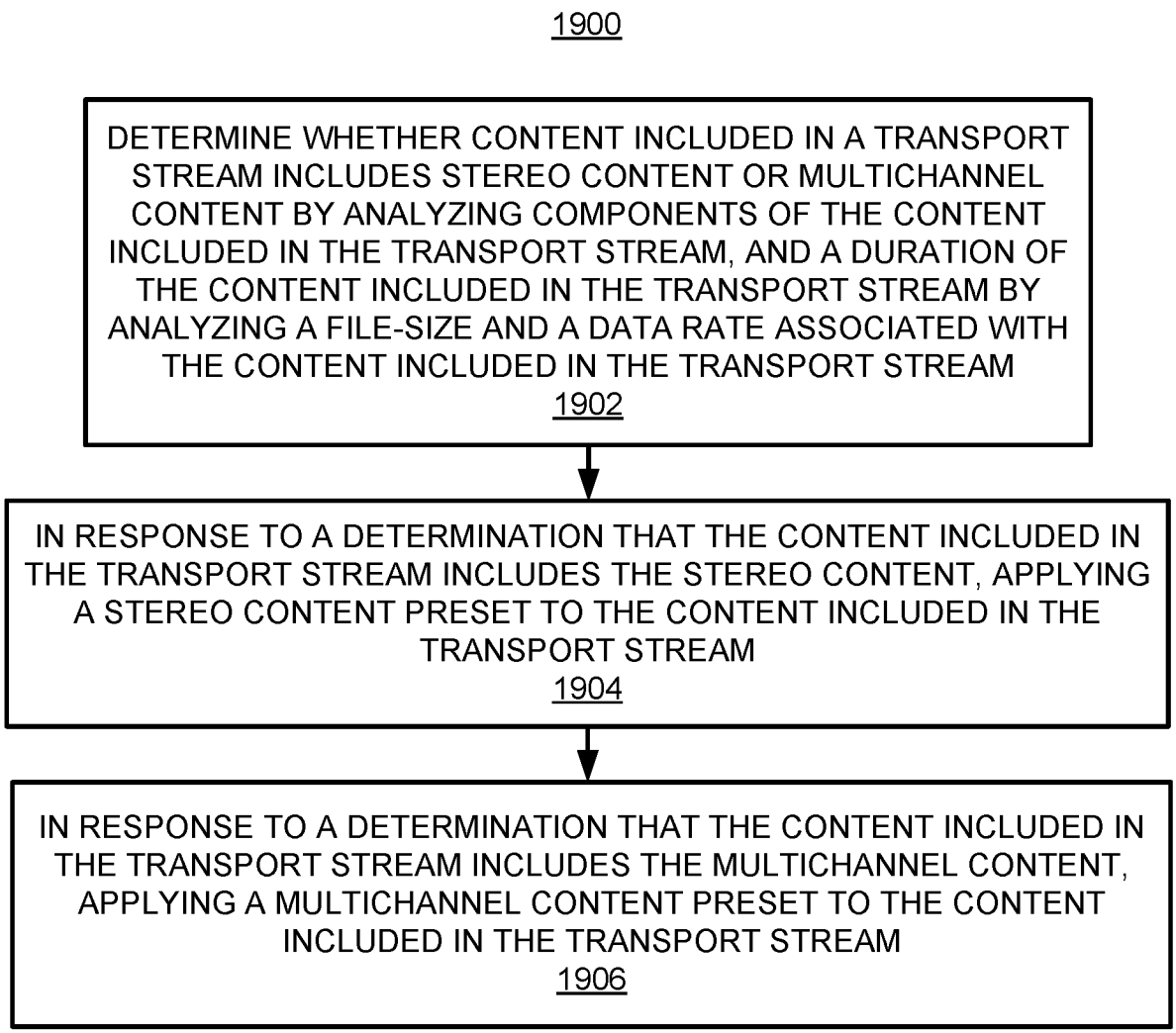

1900
DETERMINE WHETHER CONTENT INCLUDED IN A TRANSPORT STREAM INCLUDES STEREO CONTENT OR MULTICHANNEL CONTENT BY ANALYZING COMPONENTS OF THE CONTENT INCLUDED IN THE TRANSPORT STREAM, AND A DURATION OF THE CONTENT INCLUDED IN THE TRANSPORT STREAM BY ANALYZING A FILE-SIZE AND A DATA RATE ASSOCIATED WITH THE CONTENT INCLUDED IN THE TRANSPORT STREAM
1902
IN RESPONSE TO A DETERMINATION THAT THE CONTENT INCLUDED IN THE TRANSPORT STREAM INCLUDES THE STEREO CONTENT, APPLYING A STEREO CONTENT PRESET TO THE CONTENT INCLUDED IN THE TRANSPORT STREAM
1904
IN RESPONSE TO A DETERMINATION THAT THE CONTENT INCLUDED IN THE TRANSPORT STREAM INCLUDES THE MULTICHANNEL CONTENT, APPLYING A MULTICHANNEL CONTENT PRESET TO THE CONTENT INCLUDED IN THE TRANSPORT STREAM
1906

FIG. 19

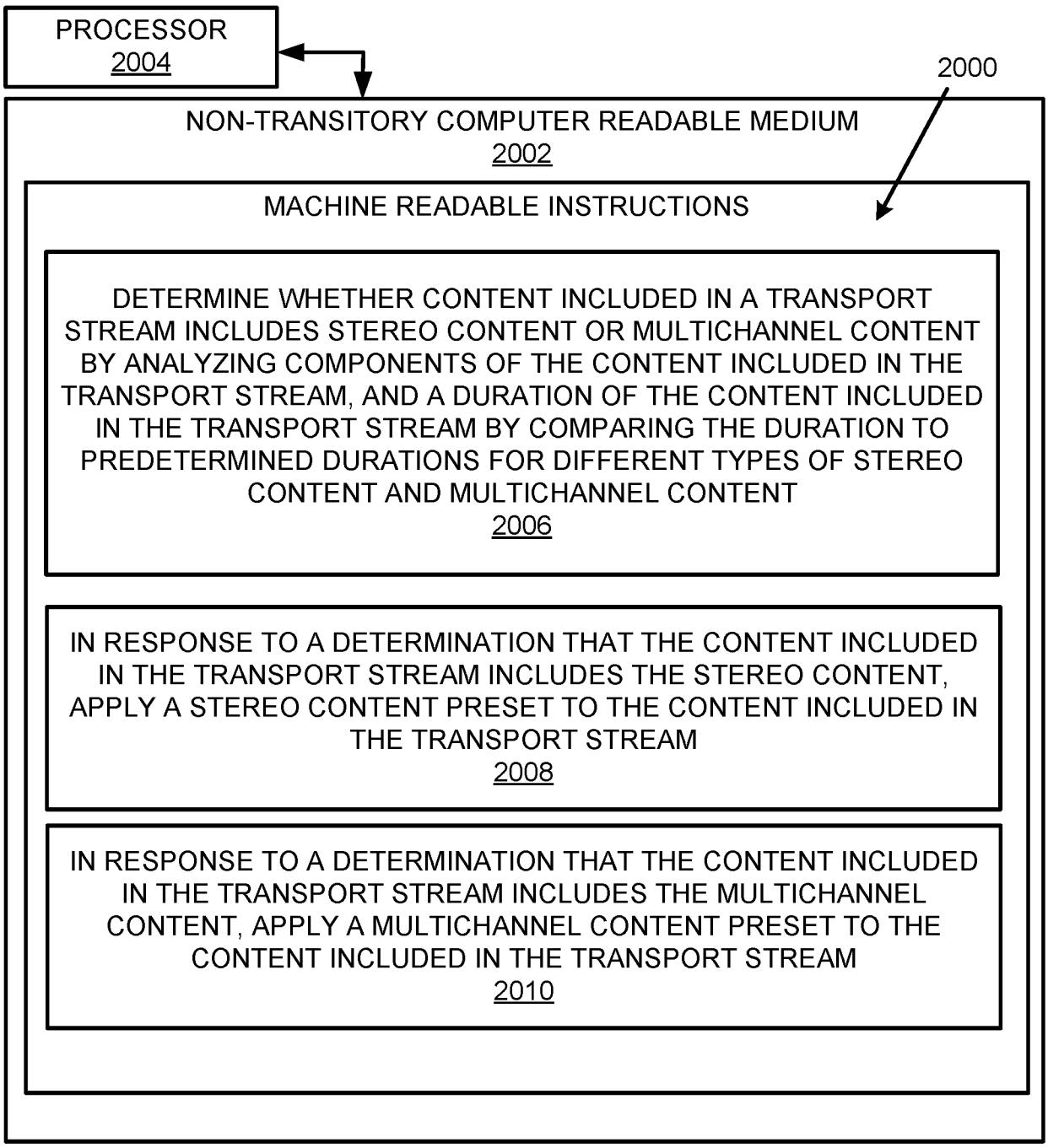

PROCESSOR
2004
2000
NON-TRANSITORY COMPUTER READABLE MEDIUM
2002
MACHINE READABLE INSTRUCTIONS
DETERMINE WHETHER CONTENT INCLUDED IN A TRANSPORT STREAM INCLUDES STEREO CONTENT OR MULTICHANNEL CONTENT BY ANALYZING COMPONENTS OF THE CONTENT INCLUDED IN THE TRANSPORT STREAM, AND A DURATION OF THE CONTENT INCLUDED IN THE TRANSPORT STREAM BY COMPARING THE DURATION TO PREDETERMINED DURATIONS FOR DIFFERENT TYPES OF STEREO CONTENT AND MULTICHANNEL CONTENT
2006
IN RESPONSE TO A DETERMINATION THAT THE CONTENT INCLUDED IN THE TRANSPORT STREAM INCLUDES THE STEREO CONTENT, APPLY A STEREO CONTENT PRESET TO THE CONTENT INCLUDED IN THE TRANSPORT STREAM
2008
IN RESPONSE TO A DETERMINATION THAT THE CONTENT INCLUDED IN THE TRANSPORT STREAM INCLUDES THE MULTICHANNEL CONTENT, APPLY A MULTICHANNEL CONTENT PRESET TO THE CONTENT INCLUDED IN THE TRANSPORT STREAM
2010

FIG. 20

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 2017/030153

A. CLASSIFICATION OF SUBJECT MATTER

H04N 21/439 *(2011.01)*
H04N 21/434 *(2011.01)*
H04H 60/56 *(2009.01)*

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

H04N 5/00-5/44, 7/00-7/18, 21/00-21/65, H04H 60/00, 60/35-60/66, G11B 27/00-27/34

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

PatSearch (RUPTO internal), USPTO, PAJ, Esp@cenet, Information Retrieval System of FIPS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2014/0095179 A1 (ELECTRONICS AND TELECOMMUNICATIONS RESEARCH INSTITUTE) 03.04.2014, paragraphs [0081], [0084]-[0086], [0097]-[0102], [0121]-[0128], clams 1-3 fig. 6-8	1, 10-13
Y		2-9, 14, 15
Y	US 2006/0203807 A1 (AI LOGIX INC) 14.09.2006, paragraphs [0309], [0367]-[0370]	3, 4
Y	US 2012/0087416 A1 (ROSS TIM) 12.04.2012, paragraphs [0035]-[0042], [0057]-[0077]	2, 5-6, 14, 15
Y	US 2012/0226494 A1 (PANASONIC CORPORATION) 06.09.2012, paragraphs [0060]-[0061], [0077]-[0095]	7-9

☐ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
28 November 2017 (28.11.2017)	07 December 2017 (07.12.2017)

Name and mailing address of the ISA/RU:
Federal Institute of Industrial Property,
Berezhkovskaya nab., 30-1, Moscow, G-59,
GSP-3, Russia, 125993
Facsimile No: (8-495) 531-63-18, (8-499) 243-33-37

Authorized officer

V. Voropai

Telephone No. (495) 531-64-81