



(51) International Patent Classification:

G06F 21/10 (2013.01) G06T 1/00 (2006.01)
H04L 9/32 (2006.01) G10L 19/018 (2013.01)

(21) International Application Number:

PCT/US2018/031912

(22) International Filing Date:

09 May 2018 (09.05.2018)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/503,833 09 May 2017 (09.05.2017) US

(71) Applicant: VERIMATRIX, INC. [US/US]; 6059 Comer-
stone Court West, San Diego, CA 92121-3713 (US).

(72) Inventor: THORWIRTH, Niels, J.; 3155 Redwood Street,
San Diego, CA 92104 (US).

(74) Agent: SUNG, Brian, K.; KPPB LLP, 2190 S. Towne Cen-
tre Place, Suite 300, Anaheim, CA 92806 (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, JO, 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.

(54) Title: SYSTEMS AND METHODS OF PREPARING MULTIPLE VIDEO STREAMS FOR ASSEMBLY WITH DIGITAL WATERMARKING

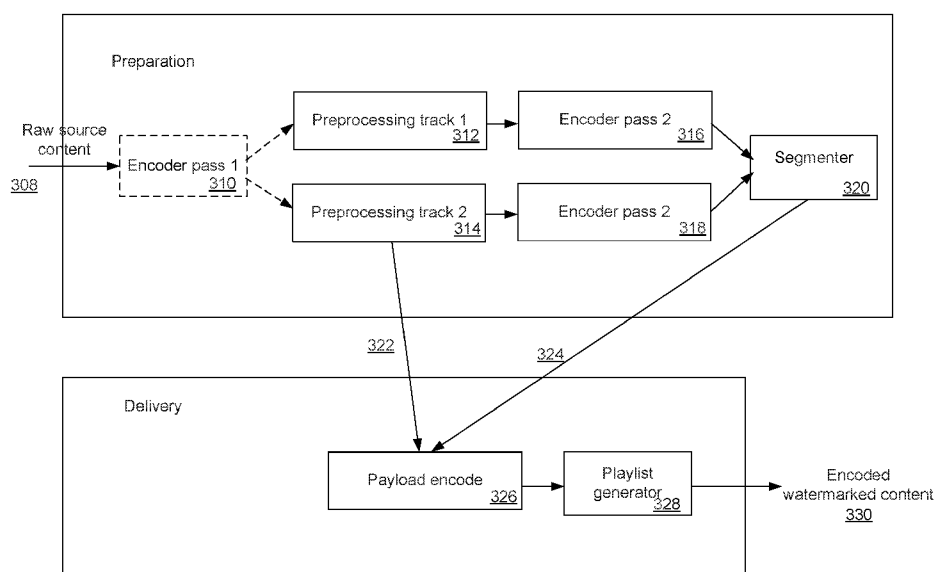


FIG. 3

(57) Abstract: Systems and methods for encoding multiple video streams with digital watermarking for adaptive bitrate streaming in accordance with embodiments of the invention are disclosed. In one embodiment, a method for preprocessing multimedia content into streams with watermark information includes receiving a source content media stream, generating at least two variant preprocessed streams for each received source content media stream, where each variant preprocessed stream includes different watermark information in the same locations as the other variant preprocessed streams and where marked locations are spaced apart at least a distance equal to a predetermined maximum segment size, generating a set of embed location information describing marked locations in the variant preprocessed streams, generating at least one variant output stream from each variant preprocessed stream using video compression, partitioning each variant output stream into a set of segments, where each segment is no longer than the predetermined maximum

(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, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

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

segment size and contains at most one copy of the watermark information, generating a set of segment boundary information describing the boundaries of segments within the variant output streams and the boundaries are the same between variant output streams, and generating a segment selection list using the set of embed location information and the set of segment boundary information, where the segment selection list includes only one variant segment for each segment according to a watermark sequence and the digits of the watermark sequence correspond to the watermark information applied to each variant preprocessed stream.

Systems and Methods of Preparing Multiple Video Streams for Assembly with Digital Watermarking

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] The present invention claims priority under 35 U.S.C. § 119(e) to U.S. Provisional Patent Application No. 62/503,833, entitled “Universal Creation and Selection Process of Adaptive Bitrate Segments to Enable Digital Watermarking”, to Thorwirth filed May 9, 2017, the disclosure of which is incorporated by reference herein in its entirety.

FIELD OF THE INVENTION

[0002] The present invention generally relates to video encoding and more specifically to systems and methods for efficiently encoding multiple streams of video content with digital watermarking for adaptive bitrate streaming from a source video stream.

BACKGROUND OF THE INVENTION

[0003] The term streaming media describes the playback of media on a playback device, where the media is stored on a server and continuously sent to the playback device over a network during playback. Typically, the playback device stores a sufficient quantity of media in a buffer at any given time during playback to prevent disruption of playback due to the playback device completing playback of all the buffered media prior to receipt of the next portion of media. Adaptive bitrate streaming (ABR), or adaptive streaming, involves detecting the present streaming conditions (e.g. the user's network bandwidth and CPU capacity) in real time and adjusting the quality of the streamed media accordingly. Typically, the source media is encoded at multiple bit rates and the playback device or client switches between streaming the different encodings depending on available resources. When a playback device commences adaptive bitrate streaming, the playback device typically starts by requesting portions of media from the lowest bitrate streams (where alternative streams are available). As the playback device downloads the requested media, the playback device can measure the

available bandwidth. In the event that there is additional bandwidth available, the playback device can switch to higher bitrate streams.

[0004] In adaptive streaming systems, the source media is typically stored on a media server as a top level index file pointing to a number of alternate streams that contain the actual video and audio data. Each stream is typically stored in one or more container files. Different adaptive streaming solutions typically utilize different index and media containers. Commonly used media container formats include Matroska, the MP4 container format specified in MPEG-4 Part 14 (i.e. ISO/IEC 14496-14), and the MPEG transport stream (TS) container specified in MPEG-2 Part 1 (i.e. ISO/IEC Standard 13818-1). The MP4 container format is utilized in ABR systems such as DASH including as specified by ISO in ISO/IEC 23009-1 Information technology — Dynamic adaptive streaming over HTTP (DASH) — Part Media presentation description and Segment formats, IIS Smooth Streaming and Flash Dynamic Streaming. The TS container is used in HTTP Adaptive Bitrate Streaming. The video in the alternate streams can be encoded pursuant to a variety of block-oriented video compression standards (or codecs) such as High Efficiency Video Coding (HEVC/H.265) specified jointly by the ISO/IEC Moving Picture Experts Group (MPEG) and the International Telecommunication Union Telecommunication Standardization Sector (ITU-T) of Geneva, Switzerland and H.264/MPEG-4 AVC (Advanced Video Coding) standard specified by the ITU-T.

SUMMARY OF THE INVENTION

[0005] Systems and methods for encoding multiple video streams with digital watermarking for adaptive bitrate streaming in accordance with embodiments of the invention are disclosed. In one embodiment, a method for preprocessing multimedia content into streams with watermark information includes receiving a source content media stream, generating at least two variant preprocessed streams for each received source content media stream, where each variant preprocessed stream includes different watermark information in the same locations as the other variant preprocessed streams and where marked locations are spaced apart at least a distance equal to a predetermined maximum segment size, generating a set of embed location information

describing marked locations in the variant preprocessed streams, generating at least one variant output stream from each variant preprocessed stream using video compression, partitioning each variant output stream into a set of segments, where each segment is no longer than the predetermined maximum segment size and contains at most one copy of the watermark information, generating a set of segment boundary information describing the boundaries of segments within the variant output streams and the boundaries are the same between variant output streams, and generating a segment selection list using the set of embed location information and the set of segment boundary information, where the segment selection list includes only one variant segment for each segment according to a watermark sequence and the digits of the watermark sequence correspond to the watermark information applied to each variant preprocessed stream.

[0006] Another embodiment includes generating a playlist from the segment selection list, where the playlist includes information enabling a playback device to retrieve watermarked content.

[0007] A further embodiment includes generating a media file from the segment selection list, where the media file includes content segments marked in a manner corresponding to the watermark sequence of the segment selection list.

[0008] Another yet embodiment includes generating at least one variant output stream from each variant preprocessed stream using video compression includes generating a set of variant alternative streams from each variant preprocessed stream, where the variant alternative streams in each set differ in bitrate and the available bitrates in each set is the same.

[0009] In another embodiment, generating at least one variant output stream from each variant preprocessed stream using video compression includes encoding only once segments that are not marked.

[0010] In still another embodiment, the watermark sequence is a binary representation of the different ways that the selected segments are marked.

[0011] In yet another embodiment, the watermark sequence is generated from a subscriber identifier associated with a user, an asset identifier associated with a piece of content, and a transaction identifier associated with a purchase transaction.

BRIEF DESCRIPTION OF THE DRAWINGS

[0012] FIG. 1 is a system diagram of an adaptive streaming system in accordance with embodiments of the invention.

[0013] FIG. 2 conceptually illustrates a media server configured to encode streams of video data for use in adaptive streaming systems in accordance with embodiments of the invention.

[0014] FIG. 3 conceptually illustrates components of an encoding process in accordance with embodiments of the invention.

[0015] FIG. 4 is a flow chart illustrating a process for encoding streams of video in accordance with embodiments of the invention.

[0016] FIG. 5 is a flow chart illustrating a process for payload encoding in generating a playlist in accordance with embodiments of the invention.

[0017] FIG. 6 conceptually illustrates examples of content during various stages of encoding in accordance with embodiments of the invention.

[0018] FIG. 7 conceptually illustrates examples of content during various stages of encoding in accordance with embodiments of the invention.

DETAILED DISCLOSURE OF THE INVENTION

[0019] Turning now to the drawings, systems and methods for encoding multiple video streams with digital watermarking for adaptive bitrate streaming in accordance with embodiments of the invention are illustrated. The present application describes several embodiments of a system that allows the preparation of content for individualized delivery. One common format of delivery is the segmentation of content into content segments or chunks in different bitrates, that allow the client to choose a bitrate that can playback without interruption given current conditions, e.g., bandwidth (also called adaptive bitrate streaming or ABR). The segments can be prepared in copies of different variations as explained in U.S. Patent Publication No. 2013/0054972, the portions of which relevant to preparing media streams is incorporated by reference in its entirety. The sequence of how segments are marked with different watermarks can be used as a unique identifying watermark sequence, as discussed in U.S. Patent

Publication No. 2011/0129116, the portions of which relevant to watermark sequences is incorporated by reference in its entirety. To create the segments the encoder may be aware of the locations that contain variations and create alternative segments during encode. This approach however may require a modification of the encoder. A simpler approach is to create different content streams, containing different variations, typically at the same locations. The encoder can take the variations and encode the streams as independent streams, without needing to be aware of the marked locations or any other information about the embedding approach.

[0020] A segmenter can partition the content into different segments to enable bitrate switching by a client device. Traditionally this is done for each bitrate at the same locations to enable seamless switching between segments of different bitrates. This same process can also be applied to all marked streams (typically 2) in all bitrates. This can also occur independent of the knowledge of the embedding location and the variations in the stream and will result in segments that each contain at most one variation if, during the preparation of the content, the locations have been spread apart with a distance larger than the maximum segment size that the segmenter will create. In this process some segments will not have a mark in some embodiments. Comparing the locations (e.g. frame numbers) of marked frames with the segments they are in (after ABR segmentation) the marks can be selected:

[0021] To now enable the selection of segments in a way to enable a unique and predetermined sequence of segments that represents a bitstring unique to a client, user or transaction, a Payload Encode module can be configured to read a list of the embedding locations (encoded as frame number or frame display time) and compare it to a list of the segment start and duration to derive the segments that contain varying segments and vary in order to set bits. The list defining the segments may be a standard defined list such as

[0022] .m3u8 - used for Apple HLS

[0023] .mpd - manifest used for MPEG DASH streaming format

[0024] SMIL - Synchronized Multimedia Integration Language

[0025] VAST -Video Ad Serving Template:

[0026] VPAID – Video Player Ad Interface Definition:

[0027] VMAP— Video Multi-Ad Playlist

[0028] or other formats, including non-standard.

[0029] The comparison process may also be used to identify the segments that are not used for embedding (i.e., corresponding segments across streams are not marked in any of the streams) and one of the copies can be removed for them since they do not vary between different tracks / embedded versions. This can reduce the amount of required storage on content servers hosting content files which may be with a CDN (content delivery network).

[0030] In adaptive streaming systems, multimedia content is encoded as a set of alternative streams of video data. Typically each alternative stream of video data is encoded using the same source multimedia content. Adaptive streaming systems are typically configured allow for streaming of multimedia content encoded at different bitrates and resolutions over a network, such as the Internet. An aim of adaptive streaming systems is to stream the highest quality version of a piece of multimedia content that can be supported based upon current streaming conditions at that particular time. Multimedia content can include video and audio data, subtitles, and/or other related metadata. In order to provide the highest quality video experience independent of the network data rate, adaptive streaming systems are configured to switch between the available sources of video data throughout the delivery of the video data according to a variety of factors, including, but not limited to, the available network data rate and video decoder performance. When streaming conditions deteriorate, an adaptive streaming system typically attempts to switch to multimedia streams encoded at lower bitrates. In the event that the available network data rate cannot support streaming of the stream encoded at the lowest bitrate, then playback is often disrupted until a sufficient amount of content can be buffered to restart playback.

[0031] To create the multiple sources of video data utilized in adaptive streaming systems, a source encoder can be configured to encode a plurality of alternative streams of video data from a source video contained in a piece of multimedia content. In accordance with several embodiments of the invention, a source encoder may be implemented using a media source and/or a media server. In accordance with some

embodiments of the invention, a set of alternative streams of video data based on the same source video may contain video data at the same resolution but differing bitrates.

Adaptive Streaming System Architecture

[0032] Adaptive streaming systems in accordance with certain embodiments of the invention are configured to generate multiple streams of video that can be made available for streaming to user devices. In many embodiments of the invention, an adaptive streaming system includes a source encoding server that performs the encoding of multiple streams of video from source media. In several embodiments, the encoding includes preprocessing an input stream to embed watermark information into different variants of the content and then video compression of the different preprocessed streams into different bitrates to create ABR playback streams. An adaptive streaming system in accordance with embodiments of the invention is illustrated in FIG. 1. The illustrated adaptive streaming system 10 includes a source encoding server 12 configured to encode source media as a number of alternative streams. The source media may be stored on the encoding server 12 or retrieved from a media source 13. As will be discussed further below, the source encoding server 12 can generate container files containing the encoded streams, at least a plurality of which are alternative streams of encoded video. In some embodiments, the encoding server makes a first pass to make certain determinations about the characteristics of the source stream and/or output stream and a second pass to encode the content into multiple output streams, where the streams can have various resolutions and/or bitrates. In some embodiments, the first pass is completed before the second pass begins. In other embodiments, the second pass can commence before the first pass is completed. In other words, computational processes for the first and second passes can run simultaneously where frames are processed by the first pass process(es) before being processed by the second pass process(es). In some embodiments, files containing output streams are uploaded to a content server 14, which can be an HTTP server or CDN. A variety of playback devices 18, 20, and 22 can then request portions of the encoded streams from the content server 14 via a network 16 such as the Internet.

[0033] Although a specific adaptive streaming system for delivering media content streams is discussed above with respect to FIG. 1, any of a variety of streaming systems can be utilized to deliver media content streams in accordance with embodiments of the invention.

Source Encoders

[0034] In the illustrated embodiment, the adaptive bitrate streaming system includes one or more source encoders capable of encoding a source streaming of video content into alternative streams of encoded video having different resolutions and/or bitrates. In many embodiments, the source encoder can be implemented using any device capable of encoding streams of multimedia, where the streams are encoded at different resolutions, sampling rates, and/or maximum bitrates. The basic architecture of an adaptive streaming system source encoder in accordance with an embodiment of the invention is illustrated in FIG. 2. The source encoder 200 includes a processor 210 in communication with memory 230 and network interface 240. In the illustrated embodiment, the volatile memory 230 includes a source encoding application 250. The processor is configured by the source encoding application 250 to encode a plurality of streams of video data from source video data 260, which is also in volatile memory. The source video data 260 may already be present in memory or may be received via the network interface 240.

[0035] In a number of embodiments, the source encoder includes multiple processors and the encoding process can be distributed among the multiple processors. In many embodiments, the source encoding application 250 can launch multiple processes that execute on one or more processors where each process is an encoder controller 280 and 290 that encodes one or more output streams. In further embodiments, each encoder controller encodes multiple output streams at the same resolution and at different bitrates. In several embodiments, an encoder controller for each of three output resolutions is launched to be executed on one or more processors.

[0036] Although a specific architecture for a source encoder is illustrated in FIG. 2, any of a variety of architectures including architectures where the video encoder 250 is located on disk or some other form of storage and is loaded into memory 230 at runtime

can be utilized to encode multimedia content in accordance with embodiments of the invention. Systems and methods for encoding of alternative streams of video data in accordance with embodiments of the invention are discussed further below.

Encoder Preprocessing Components

[0037] Components of encoding processes that may be utilized in accordance with several embodiments of the invention are illustrated in Fig. 3. These components may be implemented in software or hardware as appropriate to a particular implementation.

[0038] Raw source content 308 is received and preprocessed by introducing watermark information to generate preprocessed streams 312 and 314. Each variant of the preprocessed streams 312 and 314 differ in that each is marked in a particular way (e.g., with a particular watermark) such that each way of watermarking can be representative of a “state.” In several embodiments, this provides a binary representation (e.g., 0 or 1, A or B). Watermarks can be visible or invisible embedded information in a particular frame or across frames, as may be appropriate to any particular application. Watermark information may be inserted in certain locations that are the same in each stream. In many embodiments of the invention, preprocessing utilizes a maximum segment size so that locations that are marked in the content are spaced apart such that, when the content is segmented, each segment has at most one piece of watermark information. If a particular segment is marked, it is representative of that state (e.g., 0 or 1, A or B). The set of locations that are marked 322 are provided to payload encoder 326.

[0039] Watermarked encoded streams 316 and 318 are created by applying video compression to each variant of the preprocessed streams 312 and 314. This can include converting from frequency domain to a video format representation and associated actions such as determining frame format (I frame, P frame, B, frame), macroblock data, and/or motion vectors. In this way, there is a variant of a watermarked encoded stream for each variant of preprocessed stream.

[0040] A segmenter 320 partitions the encoded streams into segments each being no larger than the maximum segment size. The locations of segment boundaries 324 are provided to payload encoder 326. In some embodiments, a single output stream is

generated from each preprocessed stream 312 and 314. In other embodiments, sets of two or more alternative streams are generated from each preprocessed stream 312 and 314. Each set of alternative streams includes the same perceptual content and the watermark information applied to the corresponding source preprocessed stream and the alternative streams within each set are encoded at different bitrates. In this way, the sets of encoded streams may be used to provide adaptive bitrate streaming of the content to a playback device.

[0041] Payload encoder generates a list of selected segments representative of a watermark sequence. The list is generated based on a particular watermark sequence. In some embodiments, a watermark sequence can be represented in binary format (e.g., 010011) and the segments are selected based on each digit in the watermark sequence and/or whether the segment is not marked. For example in one embodiment, with a watermark sequence of 010011, the segment are marked with watermark information corresponding to a zero state, a one state, a zero state, a zero state, a one state, and a one state. In some embodiments, the segment list can be used to create a media file containing the segments. In other embodiments, the segment list can be used to create a playlist that lists information which a playback device can use to retrieve the segments of content data (e.g., addresses and/or byte ranges).

Encoding Content with Digital Watermarking for Adaptive Bitrate Streaming

[0042] A process for encoding content including watermark information in accordance with an embodiment of the invention is illustrated in FIG. 4. The encoding process 300 includes receiving media content (410). An example of a mezzanine source file in accordance with several embodiments of the invention is conceptually illustrated in Fig. 6. The content can be contained in a single input file (e.g., a multimedia file or container format file), which can be referred to as a mezzanine file, or collection of media files. The content may also be an input stream of video received by the encoder. In several embodiments of the invention, the encoding process is implemented on an encoding server as an encoding application.

[0043] In some embodiments of the invention, the encoding process includes making a first pass in processing the source media stream (412). A first pass is often intended

to make an initial analysis of the content in the source input video stream. For example, it can determine the correct allocation of bits for video streams that require variable amount of bits per frame to maintain quality, also known as variable bitrate or constant quality encoding. The first pass may determine the frame type for each frame (e.g., intra, predictive, bidirectional predictive) and other parameters used in encoding output streams from the source stream. In particular IDR (Instantaneous Decoder Refresh) information can be used to estimate the location of segment boundaries, since segmentation may only happen before IDR frames. Segment boundaries are important since one information unit such as a bit typically needs to be encoded within one segment, but may cover one or more frames. Other input allows for other optimizations such as estimating perceptual properties of the content. The IDR frame information may be produced by encoder output from a first pass encode.

[0044] The process includes generating at least two variants of preprocessed streams from the source video content and watermark information. Each variant has a variation of watermark information applied to it and may be referred to as a preprocessed stream. In many embodiments of the invention, the watermark information is embedded in locations within the content, where consecutive marked locations are spaced apart at least a distance equal to a predetermined maximum segment size, which is used to limit segment sizes when encoding the content as will be discussed further below. In several embodiments, the end of a marked portion of content is at least a length greater than the maximum segment size before the start of the next marked portion of content. Two variant preprocessed streams in accordance with several embodiments of the invention are conceptually illustrated in Fig. 6.

[0045] The preprocesses creating variant streams may run as independent processes embedding by applying deterministic rules on identification of the locations to ensure to pick the same locations in all alternative variants. There may also be one process that only reads the YUV file once and outputs two separate files or outputs only frames that differ separately. These variations depend on the tradeoff of runtime vs integration complexity where the creation of a full stream or file for each variant is the slowest process with the lowest integration complexity.

[0046] This process may be applied on YUV information that is stored in individual files for each frame or in a common file for all frames or other formats, in particular those that are readily understood by existing encoders. The encoder (on the second pass if there was a separate first pass prior) is taking the marked YUV files (at least 2) and encoding them into different files of any codec and container format.

[0047] The content is encoded (414) into output streams using video compression techniques. In embodiments including two (or more) passes, the video compression portion of encoding of output streams occurs in the second (or last) pass. In embodiments having a single pass, the video compression of output streams may include the components of the first pass described above. In some embodiments, portions of the content that are not marked are encoded only one for all the streams rather than for each stream by utilizing information concerning the locations of the watermarks. This can increase complexity of the encoder in needing to recognize and skip marked portions, but reduce the amount of encoding to be performed and storage space for the unmarked portions. In other embodiments, each preprocessed stream is processed including common portions that are not marked. This can reduce complexity of the encoder, as the encoder would not need to recognize and skip marked portions.

[0048] In several embodiments of the invention, the process includes creating multiple output streams at the same resolution but at different bitrates. In several embodiments of the invention, an encoding server has multiple processors and the encoder controllers may each run on different processors or otherwise distributed across the processors. Output streams may be encoded at a constant bit rate (CBR) or variable bit rate (VBR).

[0049] The process includes partitioning (416) the encoded streams by determining segment boundaries. In many embodiments of the invention, a predetermined maximum segment size limits the size of all segments that are generated. Similarly, the segmenter does not need to be changed to accommodate the embedding locations but can create segments according to rules as applied without watermarking. This is important because the segmentation may be applied in many different modules and may be applied differently for different ABR formats (e.g. different segment size for HLS

and DASH). Segmented encoded output streams in accordance with several embodiments of the invention are conceptually illustrated in Fig. 7.

[0050] The process includes generating (418) a representation of encoded content marked with the watermark sequence. In several embodiments, this can include selecting segments from the variants of encoded streams where the markings of each segment represent a digit in the watermark sequence. The sequence of selected segments then is indicative of the watermark sequence. In some embodiments, this can include creating an asset playlist for providing a particular playback device with information it can use to retrieve content marked in a particular way. In other embodiments, this includes generating a playback file itself containing the segments of content. In many embodiments, storage of the encoded streams can be conserved by removing segments that are not marked differently. Storage and delivery of encoded segments in accordance with several embodiments of the invention is conceptually illustrated in Fig. 7.

[0051] The embedded watermark sequence may identify a subscriber or client device using at least one piece of information such as, but not limited to, a device identifier (ID), a user identifier (ID), MAC address, IP address, login name, session information, cookie, a random generated unique ID, or similar identifier. The embedded watermark sequence may utilize additional information, such as information identifying the particular piece of content (e.g., an asset identifier), time stamp, and/or operator identifier (ID).

[0052] In some embodiments, the embedded watermark sequence is generated from a user identifier, an asset identifier, and a session identifier. Additional detail on processes for payload encoding to generate an asset playlist is provided further below.

[0053] Although a specific process for encoding multiple media content streams utilizing shared statistics is discussed above with respect to FIG. 4, any of a variety of processes can be utilized to encode multiple media content streams for adaptive bitrate streaming in accordance with embodiments of the invention.

Creating Preprocessed Stream

[0054] As discussed further above, a preprocessed stream may be utilized for embedding watermarks in the content before encoding. A process for creating a preprocessed stream in accordance with several embodiments of the invention includes the following workflow:

[0055] As Input the preprocessor is receiving:

[0056] 1. A file of a movie asset, typically in high quality format like e.g. YUV to process, including meta information about the size, resolution and frame rate.

[0057] 2. Optionally an ID for the asset that is embedded in the content, in both tracks, to identify the origin of the asset as well as the asset file. Identification of the asset may be embedded with a different technology that can be read in a first extraction step allowing extraction of information that is helpful during extraction of a second mark that contains information about the individual session. The information in the first mark may relate to how the second mark has been embedded and marking locations.

[0058] 3. The preprocessor may create two or more variant tracks or may be configured to create a single track only, in which case the information on which track to create is another input – e.g. 0 or 1. In this case, the preprocess is run once for each track.

[0059] 4. Watermark configuration information such as bitrate of watermark.

[0060] 5. Preprocessing parameters required to identify location and type of modifications, as well as primary marks and information used during the delivery step.

[0061] 6. Optionally frame type file which could be provided from a first of the compression and can help to improve the selection of marking location by allowing better estimation of the segmentation locations, that can only happen on intra frame boundaries. Better estimation can make the mark more robust and allow for higher number of bits to be embedded.

[0062] **Output**

[0063] 1. Variant tracks. A track is an encoding that contains data elements to be used to combine a payload. In a simple example a first track contains all markings representing 1 and another representing 0 to allow combination of arbitrary bitstrings combining the markings at different locations.

[0064] 2. Information used during delivery that identifies the marking locations as time location in the stream or frame number.

Payload Encoding

[0065] A process for payload encoding to generate playback information of content marked by a watermark sequence for a particular playback device is illustrated in Fig. 5. The process 500 includes generating (510) a unique transaction ID. In many embodiments, the transaction ID is generated from a subscriber ID, an asset ID, and session information. In further embodiments, additional information such as, but not limited to, a time stamp and/or operator ID, are also utilized in generating the transaction ID. An operator ID may be a unique identifier or other information that is common to all content assets distributed by the operator. In some embodiments, the transaction ID may be incremented for additional records.

[0066] The process includes generating (512) a bitstring representing the transaction ID, where the bitstring may be derived in different versions that vary with different encodings in the content over time that increase security and robustness. Some of these may include negating payload, randomization payload, encrypting the payload, combining neighboring bits with XOR to bind them together, including several bits in a segment to allow bit chains against collision, and/or negating the bit sequence for some payload numbers.

[0067] In some embodiments, the encoding and variation may be the same for several content files and the encoding of each repetition is embedded with the repetition or known as meta information about the asset during extraction. The configuration may also be identified and embedded as a mark in the content. A primary mark, being the same for all copies of this asset, may be used for that purpose.

[0068] Segments of content are selected (514) that correspond to the bits represented in the bitstring. Segments are chosen based on the number of bits they encode. For example, in some embodiments this may include a choice between segments encoding 0 and 1, but in other embodiments it may also consider segments that do not encode information or that encode several bits (e.g., for 2 bits, there can be 4 segments alternatives for each segment, with patterns of 00, 01, 10, 11). In many

embodiments, segments are selected based on the binary encoding (how the segments are watermarked) using the watermark sequence of the generated bitstring, and location information of modifications and segmentation information (start and end point of segments) derived from (e.g., .m3u8 playlist, DASH .mpd).

[0069] A marked content file, or a playlist containing information on how to retrieve the marked content file, is generated (4516) by determining a sequence of marked segments corresponding to the segments that encode the desired bitstring. This can be in any of a variety of formats, for example:

[0070] .m3u8 - used for Apple HLS,

[0071] .mpd - manifest used for MPEG DASH streaming format,

[0072] and others.

[0073] Segments concatenated to a video file in format including ISOBMFF (iso based media file format / mp4) or MPEG transport stream,

[0074] including bits in a token in an access URL to be available for selection during retrieval of segment, or

[0075] a virtual file or virtual playlist.

[0076] In some embodiments, the segments are prepared individually rather than across the entire playlist. For example, if segments are not addressed individually but have the same access link that varies in the identifier of a segment or video location only. This may be the case for a templated manifest (as they can be used in MPEG DASH) or for security reasons (where the difference between segments should not be revealed in the playlist).

[0077] To enable the process to be executed later, a session may be available that allows execution of the above process on a per segment basis during segment request – on the fly. In this case the playlist is a segment selection list that contains the segment decisions and is associated with a download or playback session with a token in the request, cookies or other means. The segment selection list is interpreted during the delivery of the content. Though the process may be executed only once per asset and stored in a token with each access request. These are 510, 512, and 514 where the token can represent the unique ID, be converted to a bitstring and converted to segment selection to accelerate and simplify the selection of the segment.

[0078] The conversions may be executed on integrated libraries or modules in other applications or on independent servers that communicate with a remote API to allow e.g. playlist manipulation (i.e. transformation) or taken evaluation and re-routing.

[0079] Upon commencement of playback by a playback device, the file can be transferred to the playback device. In some embodiments in which a playlist is used, the playlist is initially provided to the playback device. The playback device can then use the information in the playlist to request portions of the content.

[0080] Although a process for payload encoding is described above with respect to an embodiment of the invention illustrated in Fig. 5, any of a variety of processed may be utilized in accordance with embodiments of the invention.

Additional Optimizations

[0081] Reduced Overhead Storage

[0082] Create additional tracks without a mark to correct for errors or segments that should not be used, e.g., assume segment boundaries and if they are not maintained, the segments are not used accordingly and replaced with unmarked segments from the neutral (e.g. 3rd) track.

[0083] Integration Points

[0084] The preprocessor logic may be integrated in a separate module, process or machine that operates on a file, or part thereof and provides the results to the video encoder machine, module or process. This provides a high level of independence between the components. Alternatively the preprocessor may also be integrated in the encoder using an Application Programming Interface (API) between them that is passing frame data along with information about the frame from the encoder to the preprocessor to be modified. In either case, the encoder may run for several output streams – in parallel or consecutively. In the case of API this could mean that each preprocessor is configured to produce only one watermarking alternative variant stream.

[0085] Embedding Alternative

[0086] While the above description ensures spacing of the locations in minimum intervals greater than maximum segment length, an alternative approach is embedding

the locations in possibly smaller distances which may include continues embedding over time. The segmentation and embedding in this case may include several marks in one single segment. The number of marks per segment can be determined during extraction when the locations and segmentation of the content are used during extraction to determine which segments have been encoded for the stream that is being extracted. This does require the knowledge of the embedding location and segmentation during extraction either explicitly or implicitly derived from static values or by guessing.

[0087] Security Additions

[0088] Various elements to secure the execution of the playlist are outlined in U.S. Patent Publication No. 2013/0054972 to Thorwirth, the disclosure of which is hereby incorporated by reference in its entirety. The Thorwirth publication describes ways to secure token based access and combine access tokens with head end control of who and for how long the token can be accessed. As well as cryptographic means including hash and encryption to secure the token.

[0089] Live Workflow

[0090] When the system is applied to a live workflow (e.g., an event in real time), the execution speed is critical and the following optimizations, while not limited to live content are particularly useful in that case:

[0091] The preprocessing of the content may be performed in a single pass, not using or creating first pass information. Embedding locations in time, in this case are chosen within a segment by using signaling for IDR frames that, in particular when combined with segment duration and fixed segment duration, indicate a new segment. Also fixed embedding locations that have the same temporal distance as the fixed segment duration enable marking that is isolated between all segments and can therefore be addressed individually when choosing individual segments. Embedding a mark in every segment has the additional advantage that the information about the segments that contain a mark is static if every segment contains a mark. The segment selection, in this case, only needs to apply the segment selection without considering unmarked segments and requiring an information stream from the preprocessor.

[0092] To allow for faster processing, the preprocessor may be integrated in the encoder in this case, using an interface that e.g. exchanges baseband frames. As in the non-live / video-on-demand case, the encoder may run in two (or more) independent instances that each create a different variant of the segments by using a preprocessor that is configured differently but without having to be otherwise adjusted to the process.

[0093] Collusion Prevention

[0094] Instead of selecting or re-routing segments, individual segments may also be created for individual sessions, e.g. including the transaction or client number in a single frame or segment. In an attack that would mix segments from differently marked streams to obfuscate the bitstring, this single segment would reveal the identity. While this can be slower to prepare a unique segment, this would not have to be done often and can be made anticipating the request (e.g. when movie starts for segment 5 minutes in) and allow time for preparation.

[0095] Between these alternatives of embedding all or many bits of information in a single segment and the approach to embed a single bit per segment, other encoding variations allow improved ability for reading payload from one or all contributing colliding streams by embedding several bits in a segment. E.g., two bits can be embedded by either allowing 4 alternative variants with 4 alternative segments in the same location for some segment or by creating a single alternative segment in 4 different segment locations, each of these alternatives representing a combination of the two bit values while the other has no marking. When used three of the four locations are used without marking while one of alternative is chosen to embed the two bits it contains. This allows a possibility of only having a maximum of one alternative segment and still embed several bits together. This can be used to limit the amount of segments that are used that contain a modification and allows embedding of several bits from one payload to facilitate the decoding and assigning related bits to one of the payload bit strings.

Applications

[0096] The systems and techniques described above can be used in a virtually limitless range of applications. The following are provided as a series of examples of applications in which systems and processes in accordance with embodiments of the invention can be useful.

[0097] Tracking Unauthorized Distribution of Secret or Copyrighted Information

[0098] A problem that faces many industries is the unauthorized distribution of information. Systems and processes in accordance with embodiments of the present invention can be used to embed marks in media information at the time of reception or display of the media. Each distributed copy can be uniquely marked with information such as a recipient identification number and a time stamp and, if the copy is publicly available or in the possession of an entity or individual that is not authorized to possess the information, the information can be uncovered and the entity or person that is the recipient of the media and the likely source of the unauthorized distribution can be identified.

[0099] In many instances, the secret or copyrighted information is passed between several different entities and/or individuals during production and authorized distribution. In several embodiments, the point from which the information was distributed without authorization can be ascertained by embedding a mark associated with the last recipient of the information prior to delivery or display. The entity or individual that is responsible for the unauthorized distribution can then be identified based upon the last mark added to the media.

[00100] A common instance, in which copyrighted information is communicated, is the distribution of copyrighted media via a network to a media player. In many embodiments, the player is a consumer electronics device such as a set top box or a personal computer. The copyrighted media is typically distributed to the player in a compressed and encrypted form. A mark can be embedded in the media in accordance with embodiments of the invention. The mark can contain information relating to the owner of the player and information identifying the time of transmission or playback. If the recipient of the information is known, the information to be embedded can be generated by the server (or head end) providing the media. The embedded information can also be stored by the server in a database that contains additional information about the transaction, such as the user's billing information and details about the receiving device. In other embodiments, the player maintains information such as player

identification number and time, which is embedded as a mark during storage and/or playback.

[00101] Another instance in which unauthorized distribution is a common problem is in the production of media. During production, content is particularly vulnerable to unauthorized distribution that can cause considerable damage to the producer of the media. In many embodiments, marks are embedded in the media during various stages of production that identify the recipient of the media and the time of the receipt of the media. If the copy is made publicly available, the mark can be uncovered and the responsible person or entity can be identified.

[00102] In many embodiments of the invention, a perceptual model is created for the distributed media and stored or distributed with the media. The perceptual model can then be used as marks are embedded in the media. In other embodiments, a perceptual model is created every time a mark is embedded in the media.

[00103] Proof of Ownership

[00104] Once media is subject to public distribution, proving ownership of copyright in the media can be problematic. In many embodiments, marks are embedded in the media to prove ownership of the content. In a number of embodiments, the mark includes ownership information or identification for copyrighted content. This mark can be automatically read and its presence or absence can be evaluated to restrict or allow distribution or to verify that a source distributing the media is a legitimate distributor of the media. It can also be used, so that entities or individuals interested in obtaining rights with respect to the content can identify the owner of the content by uncovering the embedded mark. Another use for a mark identifying the owner is to automatically identify content in publicly available database, e.g., accessible through the internet for content owned by a specific owner or group of owners to prevent further distribution. Finally this mark can be used to prove ownership in case of a dispute.

[00105] Storing Robust Meta-Information

[00106] During long term storage and archival, information stored with media can be lost either because it is deleted accidentally or because it cannot be read. Many embodiments of the invention are configured to store information concerning media

using marks. Storing information using marks can enable the information to be retrieved after format change and does not require additional storage space.

[00107] Copy Control

[00108] In many embodiments, a strong mark that is machine readable can be embedded into media. The mark can then be used by a player to control playback and/or recording of the media. The player can uncover the mark and ascertain the authorizations required to play the media. If the player has the appropriate authorizations, then the player can play the media or refuse playback otherwise.

[00109] Broadcast Monitoring

[00110] Machine readable marks in accordance with embodiments of the present invention can also be embedded in media broadcast by a television or radio station. The machine readable marks can then be used by receiver devices to automatically record the content that was broadcast and to track the frequency and time of broadcast. The mark embedded in the media can be distinct for each work and broadcast.

[00111] Secret Communication

[00112] As mentioned above, marks can be used to transmit secret information using media. While the transmission of the media can be observed the fact that the media contain an additional mark is not obvious and can be used to transmit information without allowing others to observe that information is transmitted other than the media content.

[00113] Identification of Publicly Displayed Media

[00114] Media that are publicly displayed are frequently the subject of rerecording. In many embodiments, a mark is embedded in publicly displayed media that identifies the time and/or place of public display. Should the media be rerecorded during the public display, then the mark is embedded in the unauthorized rerecording and uncovering this information can be helpful in preventing future rerecording of publicly displayed media at the venue in which the rerecording was made.

[00115] Although the present invention has been described in certain specific aspects, many additional modifications and variations would be apparent to those skilled in the art. It is therefore to be understood that the present invention may be practiced otherwise than specifically described, including various changes in the implementation

such as utilizing encoders and decoders that support features beyond those specified within a particular standard with which they comply, without departing from the scope and spirit of the present invention. Thus, embodiments of the present invention should be considered in all respects as illustrative and not restrictive.

WHAT IS CLAIMED IS:

1. A method for preprocessing multimedia content into streams with watermark information, the method comprising:

receiving a source content media stream;

generating at least two variant preprocessed streams for each received source content media stream, where each variant preprocessed stream includes different watermark information in the same locations as the other variant preprocessed streams and where marked locations are spaced apart at least a distance equal to a predetermined maximum segment size;

generating a set of embed location information describing marked locations in the variant preprocessed streams;

generating at least one variant output stream from each variant preprocessed stream using video compression;

partitioning each variant output stream into a set of segments, where each segment is no longer than the predetermined maximum segment size and is marked with watermark information in at most one location;

generating a set of segment boundary information describing the boundaries of segments within the variant output streams and the boundaries are the same between variant output streams; and

generating a segment selection list using the set of embed location information and the set of segment boundary information, where the segment selection list includes only one variant segment for each segment according to a watermark sequence and the digits of the watermark sequence correspond to the watermark information applied to each variant preprocessed stream.

2. The process of claim 1, further comprising generating a playlist from the segment selection list, where the playlist includes information enabling a playback device to retrieve watermarked content.

3. The process of claim 1, further comprising generating a media file from the segment selection list, where the media file includes content segments marked in a manner corresponding to the watermark sequence of the segment selection list.
4. The process of claim 1, wherein generating at least one variant output stream from each variant preprocessed stream using video compression includes generating a set of alternative streams from each variant preprocessed stream, where the alternative streams in each set differ in bitrate and the available bitrates in each set is the same.
5. The process of claim 1, wherein generating at least one variant output stream from each variant preprocessed stream using video compression includes compressing segments that are not marked only once.
6. The process of claim 1, wherein the watermark sequence is a binary representation of the different ways that the selected segments are marked.
7. The process of claim 1, wherein the watermark sequence is generated from a transaction identifier associated with a purchase transaction.
8. The process of claim 7, wherein the watermark sequence is also generated using a subscriber identifier associated with a user.

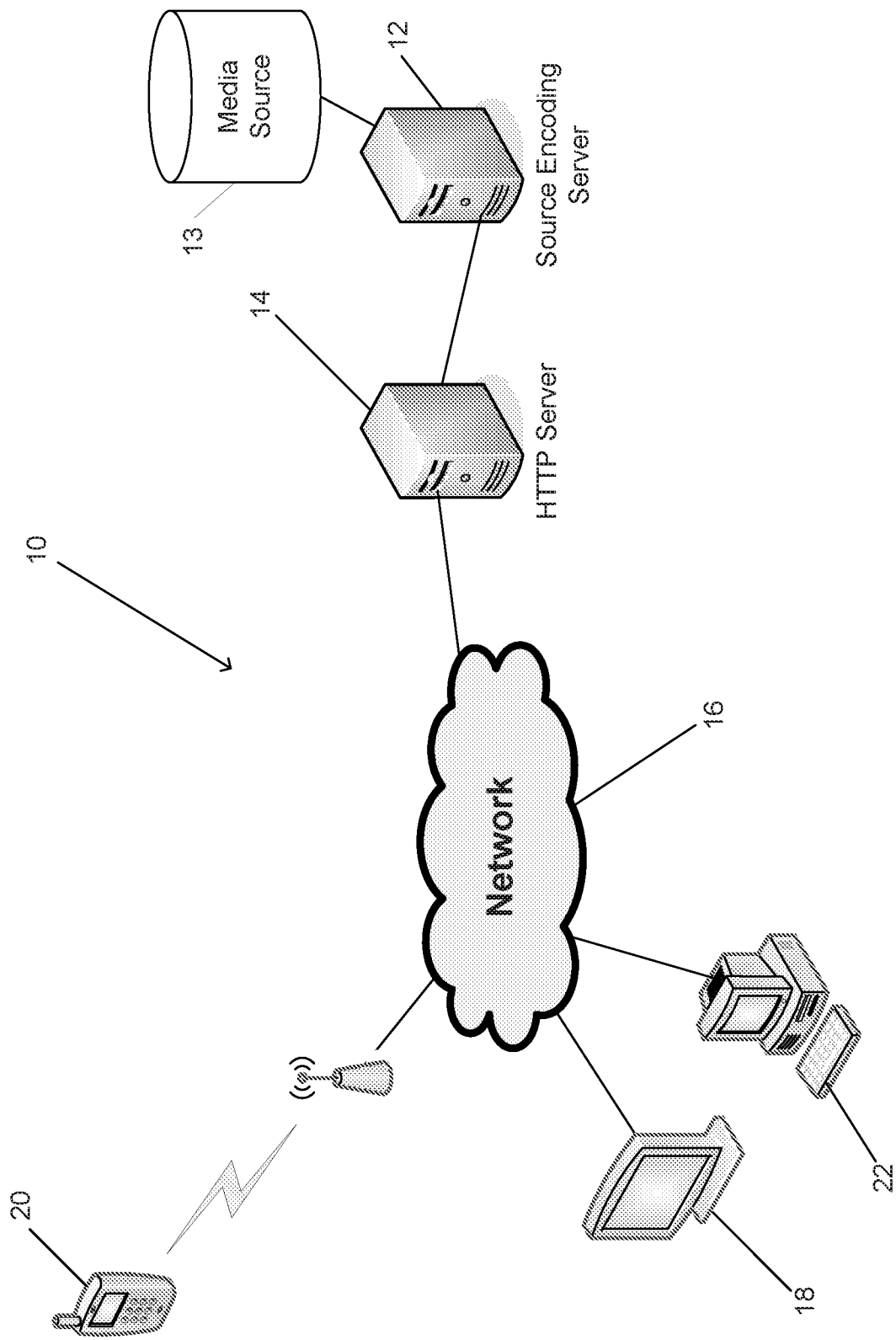


FIG. 1

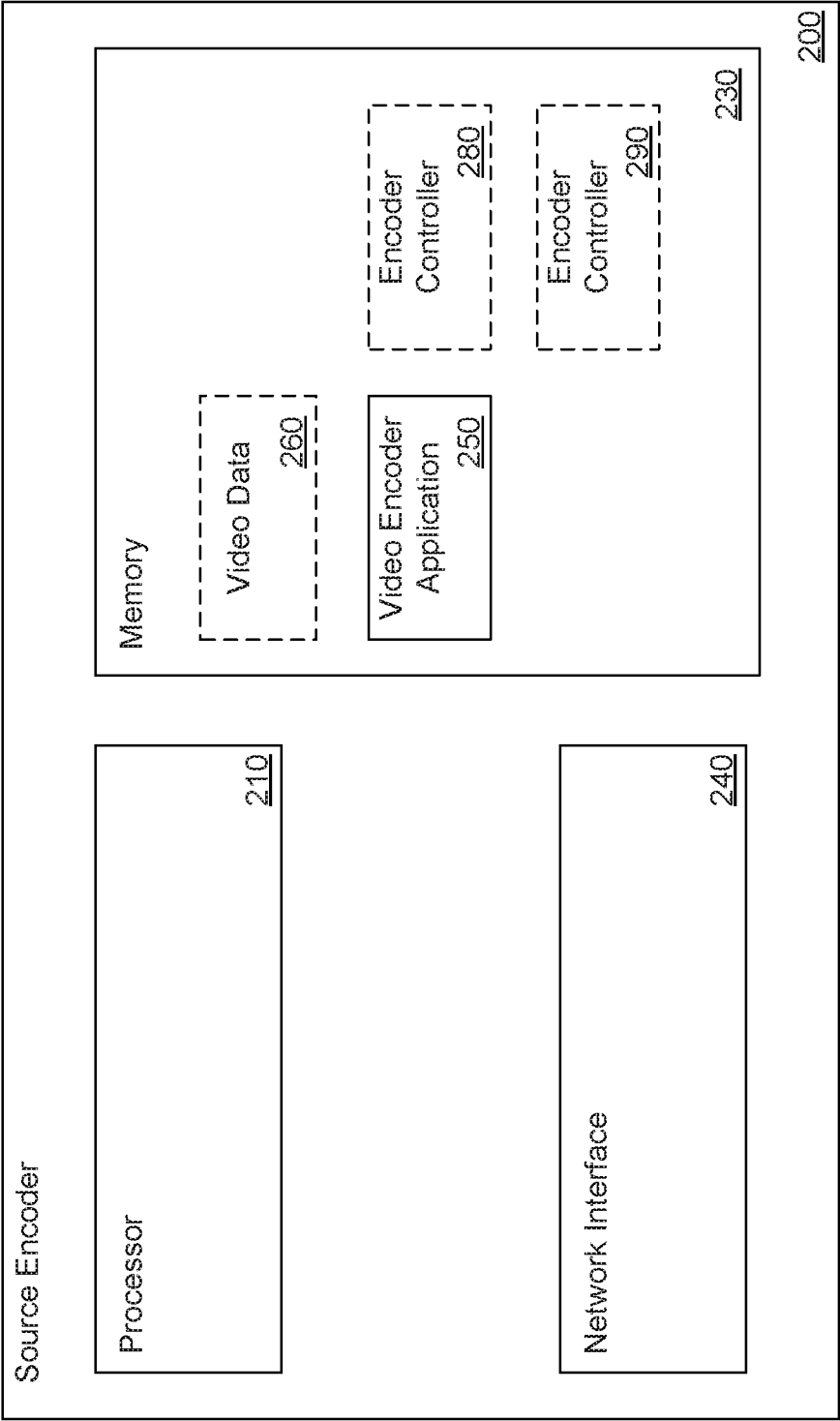


FIG. 2

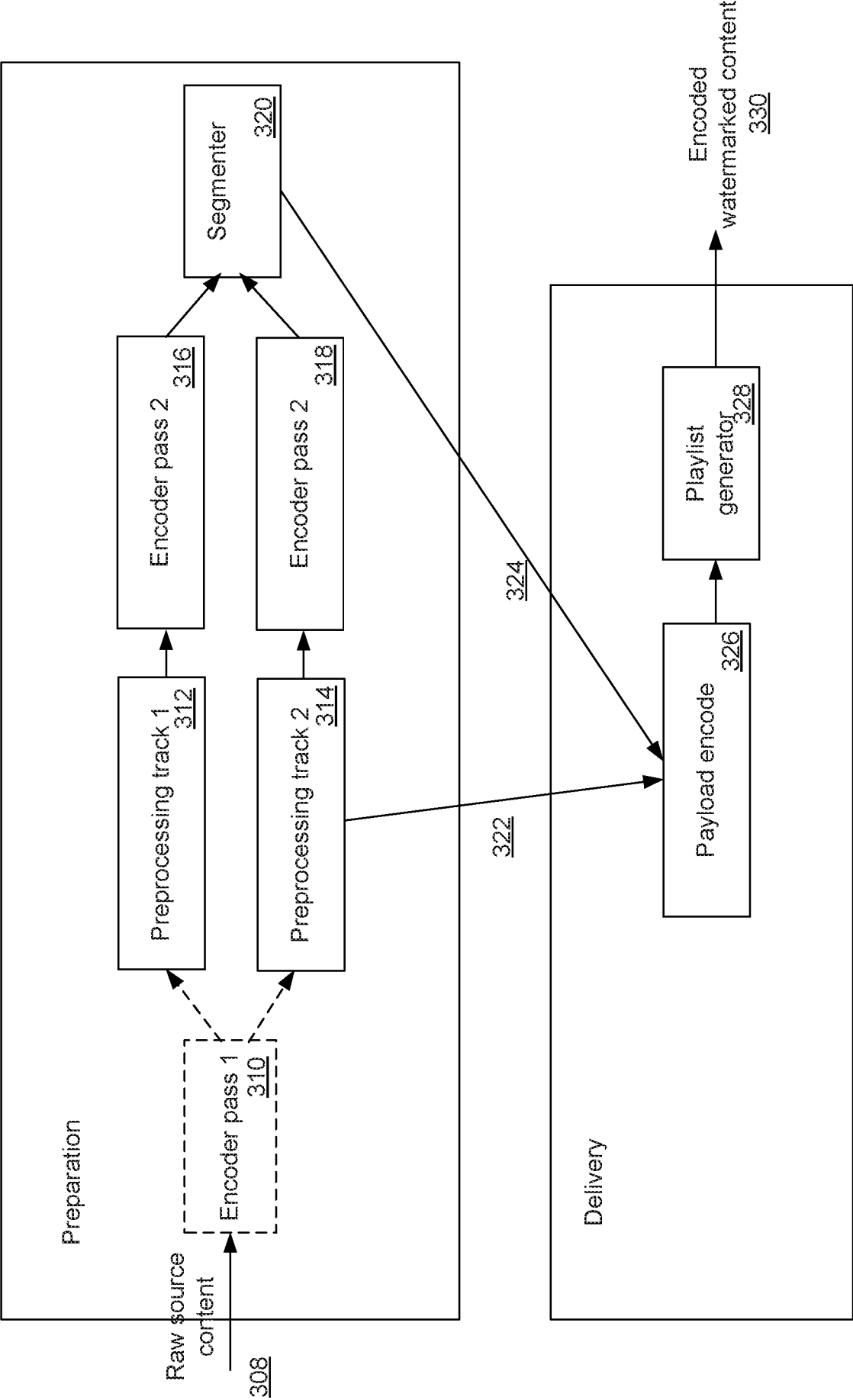
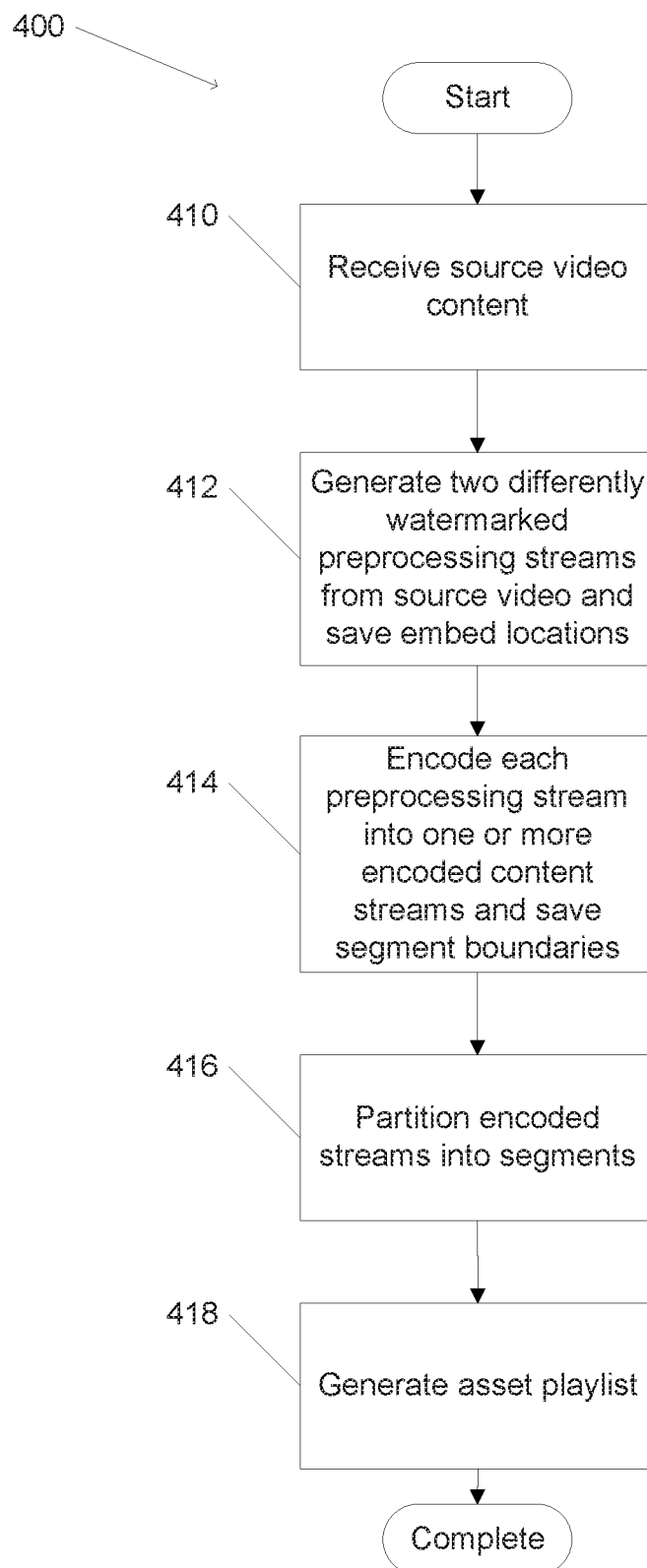
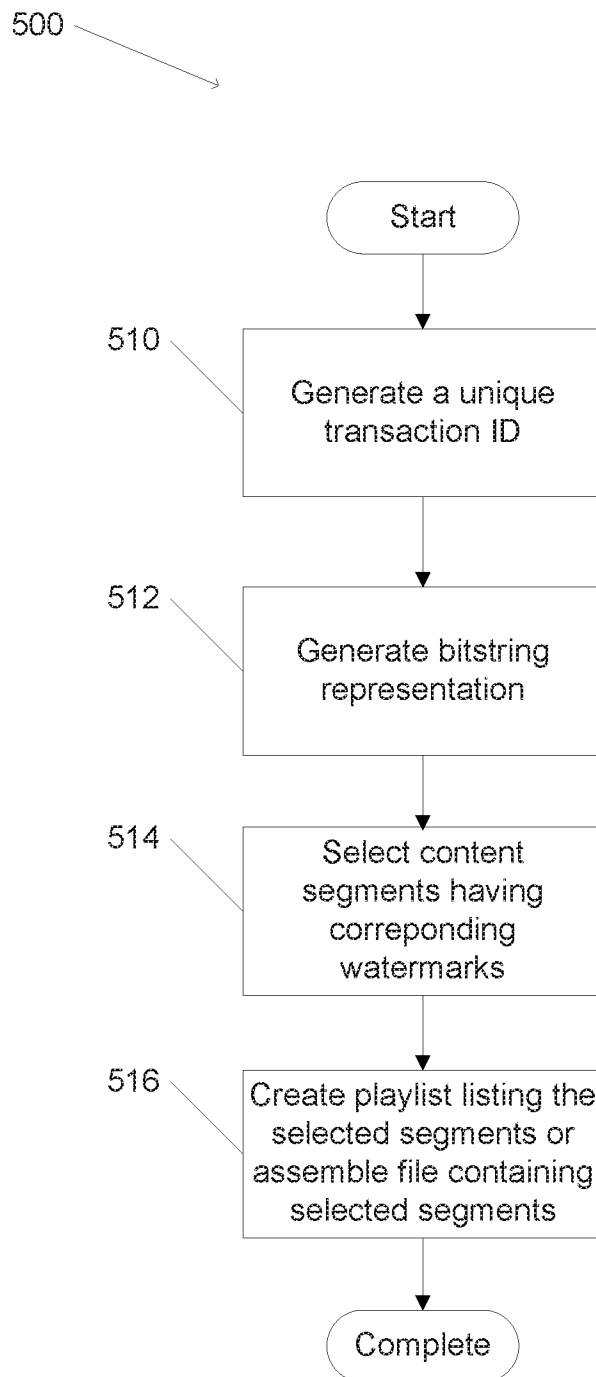


FIG. 3

4/7

**FIG. 4**

5/7

**FIG. 5**

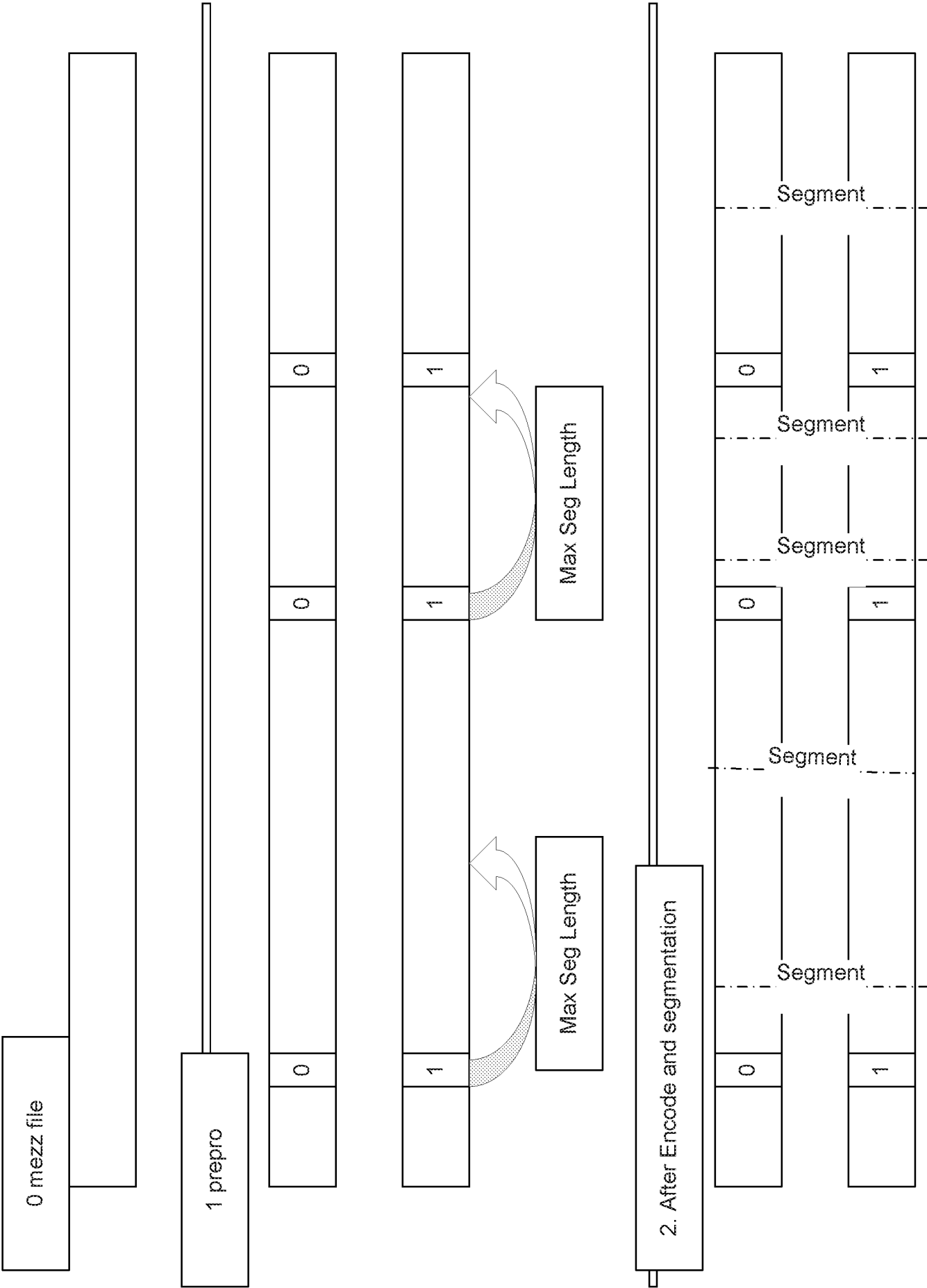


FIG. 6

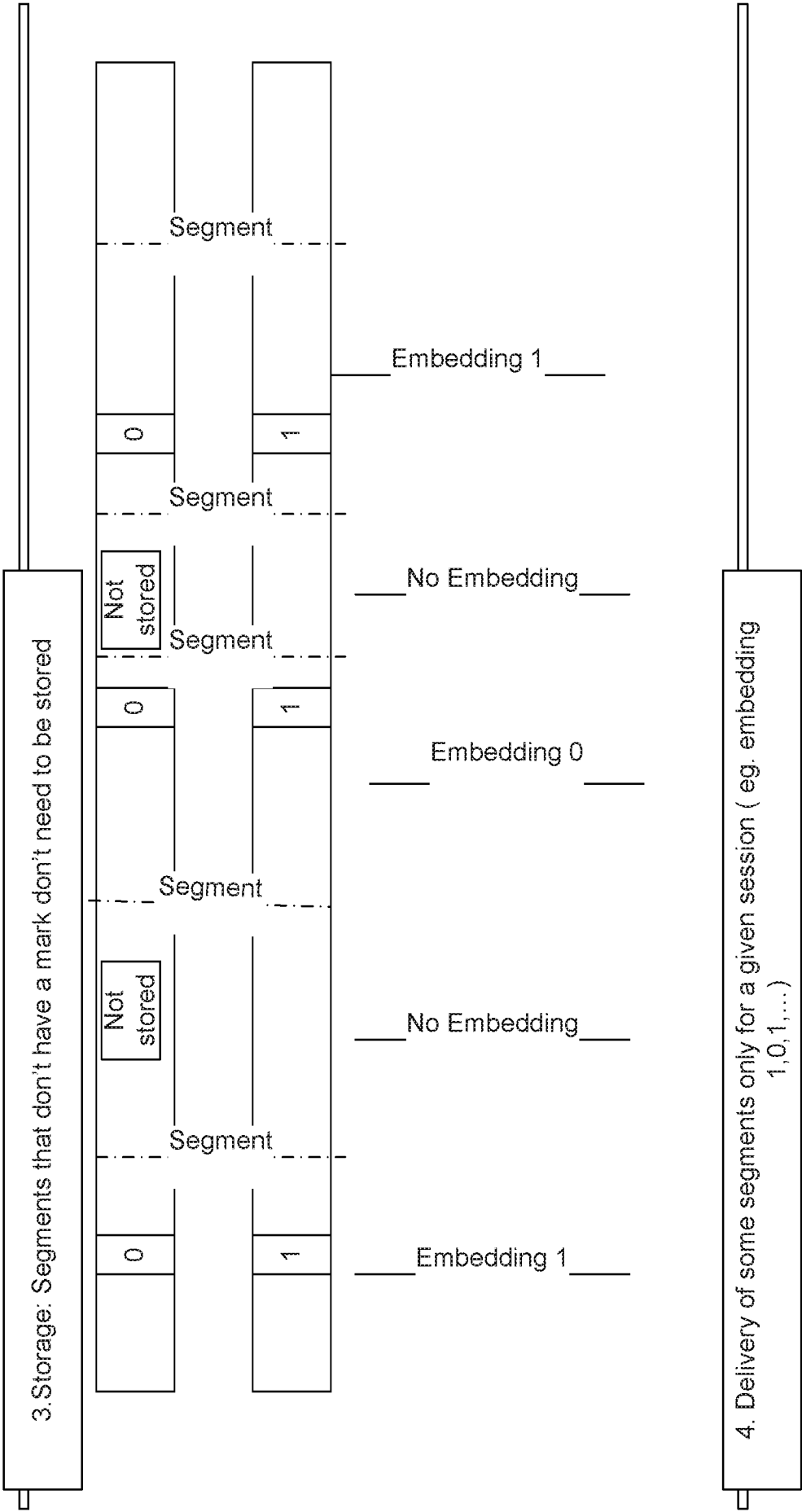


FIG. 7

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2018/031912

A. CLASSIFICATION OF SUBJECT MATTER

IPC(8) - G06F 21/10; H04L 9/32; G06T 1/00; G10L 19/018 (2018.01)

CPC - G06T 1/0021; G06F 21/10; G06T 1/0028; G06T 1/0071 (2018.05)

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)

See Search History document

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

USPC - 382/100; 382/232; 726/26 (keyword delimited)

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

See Search History document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2013/0142382 A1 (VERANCE CORPORATION) 06 June 2013 (06.06.2013) entire document	1-8
A	US 2004/0117427 A1 (ALLEN et al) 17 June 2004 (17.06.2004) entire document	1-8
A	US 2001/0051996 A1 (COOPER et al) 13 December 2001 (13.12.2001) entire document	1-8
A	US 2004/0139047 A1 (HINDLE et al) 15 July 2004 (15.07.2004) entire document	1-8
A	US 2014/0270337 A1 (VERANCE CORPORATION) 18 September 2014 (18.09.2014) entire document	1-8



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 application or patent 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

03 July 2018

Date of mailing of the international search report

23 JUL 2018

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents

P.O. Box 1450, Alexandria, VA 22313-1450

Facsimile No. 571-273-8300

Authorized officer

Blaine R. Copenheaver

PCT Helpdesk: 571-272-4300

PCT OSP: 571-272-7774