

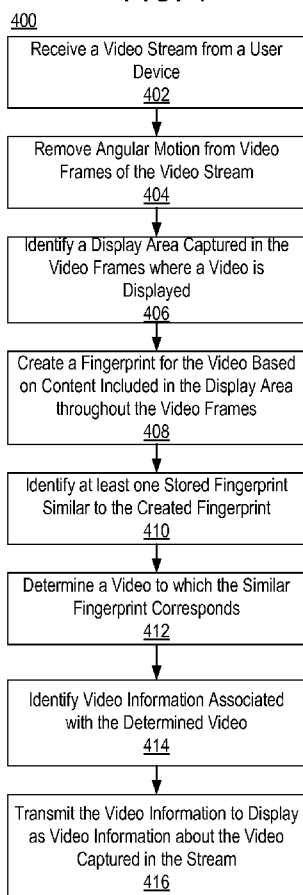


- (51) **International Patent Classification:**
H04N 21/234 (2011.01) *G06K 9/00* (2006.01)
H04N 21/4223 (2011.01) *G06F 17/30* (2006.01)
H04N 21/278 (2011.01)
- (21) **International Application Number:**
PCT/US2015/044663
- (22) **International Filing Date:**
11 August 2015 (11.08.2015)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
14/489,402 17 September 2014 (17.09.2014) US
- (71) **Applicant:** **GOOGLE INC.** [US/US]; 1600 Amphitheatre Parkway, Mountain View, CA 94043 (US).
- (72) **Inventors:** **WEICH, Tomas**; C/o Google Inc., 1600 Amphitheatre Parkway, Mountain View, CA 94043 (US). **ERB, David**; C/o Google Inc., 1600 Amphitheatre Parkway, Mountain View, CA 94043 (US).
- (74) **Agents:** **HOFFMAN, Brian, M.** et al.; Fenwick & West LLP, Silicon Valley Center, 801 California Street, Mountain View, CA 94041 (US).
- (81) **Designated States** (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR,

[Continued on next page]

(54) **Title:** IDENTIFICATION OF CAPTURED VIDEOS

FIG. 4



(57) **Abstract:** An identification system receives a request from a user device to identify a video captured in a video stream. The identification system identifies a display area captured in the video stream in which the video is displayed. The identification system identifies content include in the display area throughout the stream. Based on the identified content, the identification system creates a fingerprint for the video. The identification system identifies from a plurality of stored fingerprints at least one fingerprint similar to the created fingerprint. The identification system identifies video information associated with the similar fingerprint. The video information is transmitted to the user device and displayed as being video information about the video captured in the stream.



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.

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).

(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,

Published:

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

IDENTIFICATION OF CAPTURED VIDEOS

BACKGROUND

FIELD

[0001] Described embodiments pertain in general to videos systems, and in particular to the identification of videos.

DESCRIPTION OF THE RELATED ART

[0002] On a daily basis media is presented to people everywhere they go. For example, when a person goes to a shopping mall, the person may see video ads for new products, movie trailers, music videos, etc. If a person sees a video that is playing and is interested in it, the person may want to obtain certain information about the video, such as the name of the video and where it can be found. However, in certain circumstances obtaining such information may be difficult. Continuing with the shopping mall example from above, assume the person is at a store and passes a television display that is showing an interesting movie. The person would like to rent the movie and needs to know the name of the movie but the name is nowhere to be found. The person could ask other shoppers at the store or an employee of the store for the name of the movie. However, asking someone may be embarrassing to the person, may take a long time (e.g., searching for an employee at the store) and whoever the person asks may not know the name of the movie.

SUMMARY

[0003] Described embodiments provide methods, computer program products, and systems for identifying videos captured in video streams. When a user is watching a video and would like more information about the video (e.g., the name of the video and information for accessing the video), the user uses a user device 102 to create a video stream that captures playback of a portion of the video. The user device 102 transmits the video stream to an identification system and requests identification of the video captured in the video stream.

[0004] The identification system identifies a display area captured in the video stream in which the video is displayed. For example, the identified display area corresponds to an object (e.g., display of a television) in which the video was displayed during the creation of the video stream. The identification system identifies content included in the display area throughout the stream. The identified content includes content/scenes of the video captured in stream. Based on the identified content, the identification system creates a fingerprint for the video.

[0005] The identification system identifies from a plurality of stored fingerprints at least one fingerprint similar to the created fingerprint. The identification system identifies video information associated with the similar fingerprint. The video information may include, for example, a video identifier and a link for accessing a video. The video information is transmitted to the user to describe the video captured in the stream.

[0006] The features and advantages described in this summary and the following detailed description are not all-inclusive. Many additional features and advantages will be apparent to one of ordinary skill in the art in view of the drawings, specification, and claims hereof.

BRIEF DESCRIPTION OF THE DRAWINGS

[0007] FIG. 1 is a block diagram of a video identification environment according to one embodiment.

[0008] FIG. 2 is a block diagram illustrating a functional view of a typical computer system for use as one of the entities illustrated in the environment of FIG. 1 according to an embodiment.

[0009] FIG. 3 is a block diagram illustrating modules operating within an identification system according to one embodiment.

[0010] FIG. 4 is a flow chart illustrating operations of an identification system in identifying a video captured in a stream according to one embodiment.

[0011] The figures depict various embodiments for purposes of illustration only. One skilled in the art will readily recognize from the following discussion that alternative embodiments of the structures and methods illustrated herein may be employed without departing from the principles of the embodiments described herein.

DETAILED DESCRIPTION

[0012] FIG. 1 is a block diagram of a video identification environment 100 according to one embodiment. FIG. 1 illustrates user devices 102A and 102B and identification system 104 connected via a network 106. Although a select number of each entity are shown in FIG. 1, embodiments can have more or less of each entity (e.g., additional user devices 102).

[0013] FIG. 1 and the other figures use like reference numerals to identify like elements. A letter after a reference numeral, such as “102A,” indicates that the text refers specifically to the element having that particular reference numeral. A reference numeral in the text without a following letter, such as “102,” refers to any or all of the elements in the figures bearing that reference numeral.

[0014] A user device 102 is a device used by a user to request identification of a video captured in a video stream. In one embodiment, a user device 102 includes functionality for creating video streams of objects. A video stream is video content that is comprised of multiple video frames. A user device 102 may be any device with video capturing capabilities, such as a mobile phone, a tablet computer, or a wearable computing device (e.g., glasses or a watch). When a user sees a video being displayed and the user would like more information about the video (e.g., title of the video, author, where the video can be accessed), the user uses a user device 102 to create a video stream of the video. The video stream includes multiple sequenced video frames and captures playback of a portion of the video (captures a sample of the video). In one embodiment, the video stream is 4.5 seconds in length. The video whose playback is captured by the video stream may be, for example, a video provided by a video sharing system that hosts videos uploaded by users. The video may also be a television program or a movie.

[0015] In addition to the video stream capturing playback of the video, the stream may also capture other objects that surround the display of the video. For example, if the user is in a living room and records a video being displayed on a television, in addition to the stream created including the television, the stream may also include a chair and a couch that are within the user device's 102 field of view. The user device 102 transmits the video stream to the identification system 104 and requests identification of the video.

[0016] In one embodiment, the user device 102 streams video stream to the identification system 104 as it is being created. In another embodiment, the user device 102 transmits the video stream to the identification system 104 once creation of the stream is complete (e.g., when user requests to stop recording).

[0017] In one embodiment, instead of the user device 102 creating the video stream, the video stream is created by another device (e.g., a video camera) separate from the user device 102. The video stream is transferred to the user device 102 and the user device 102 transmits the request for identification to the identification system 104 along with the video stream.

[0018] The identification system 104 is a computer system that identifies videos captured in video streams. When the identification system 104 receives a video stream with a request for a video captured in the stream to be identified, the identification system 104 identifies a display area captured in the stream in which the video is displayed. The identified area may, for example, correspond to a television display in which the video is displayed.

[0019] The identification system 104 identifies content included in the display area throughout the multiple frames of the stream. Based on the content, the identification system creates a fingerprint for the portion of the video that was captured by video stream. The identification system 104 compares the created fingerprint to stored video fingerprints. The identification system 104 identifies at least one stored fingerprint that is similar to the created fingerprint of the video captured in the stream.

[0020] The identification system 104 determines a video to which the identified stored fingerprint corresponds and retrieves video information associated with the determined video (i.e., retrieve video information associated with the identified fingerprint), including an identifier of the video. The identification system 104 provides the retrieved information to the user device 102. The information is displayed to the user of the device 102 as being information about the video captured in the stream.

[0021] The network 106 represents the communication pathways between the user devices 102 and the identification system 104. In one embodiment, the network 106 is the Internet and uses standard communications technologies and/or protocols. Thus, the network 106 can include links using technologies such as Ethernet, 802.11, worldwide interoperability for microwave access (WiMAX), 3G, Long Term Evolution (LTE), digital subscriber line (DSL), asynchronous transfer mode (ATM), InfiniBand, PCI Express Advanced Switching, etc. Similarly, the networking protocols used on the network 106 can include multiprotocol label switching (MPLS), the transmission control protocol/Internet protocol (TCP/IP), the User Datagram Protocol (UDP), the hypertext transport protocol (HTTP), the simple mail transfer protocol (SMTP), the file transfer protocol (FTP), etc.

[0022] The data exchanged over the network 106 can be represented using technologies and/or formats including the hypertext markup language (HTML), the extensible markup language (XML), etc. In addition, all or some of links can be encrypted using conventional encryption technologies such as secure sockets layer (SSL), transport layer security (TLS), virtual private networks (VPNs), Internet Protocol security (IPsec), etc. In another embodiment, the entities can use custom and/or dedicated data communications technologies instead of, or in addition to, the ones described above.

[0023] FIG. 2 is a block diagram illustrating a functional view of a typical computer system for use as one of the entities illustrated in the environment 100 of FIG. 1 according to an embodiment. Illustrated are at least one processor 202 coupled to a chipset 204. Also coupled to the chipset 204 are a memory 206, a storage device 208, a keyboard 210, a

graphics adapter 212, a pointing device 214, a network adapter 216, and a video recording device 224. A display 218 is coupled to the graphics adapter 212. In one embodiment, the functionality of the chipset 204 is provided by a memory controller hub 220 and an I/O controller hub 222. In another embodiment, the memory 206 is coupled directly to the processor 202 instead of the chipset 204.

[0024] The storage device 208 is a non-transitory computer-readable storage medium, such as a hard drive, compact disk read-only memory (CD-ROM), DVD, or a solid-state memory device. The memory 206 holds instructions and data used by the processor 202. The video recording device 224 records objects that are within a field of view of the device 224 and creates video streams that include the recorded content. The pointing device 214 may be a mouse, track ball, or other type of pointing device, and is used in combination with the keyboard 210 to input data into the computer system 200. The graphics adapter 212 displays images and other information on the display 218. The network adapter 216 couples the computer system 200 to the network 106.

[0025] As is known in the art, a computer 200 can have different and/or other components than those shown in FIG. 2. In addition, the computer 200 can lack certain illustrated components. For example, a computer system 200 acting as the identification system 104 may lack a keyboard 210, a pointing device 214, and a video stream device 224. Moreover, the storage device 208 can be local and/or remote from the computer 200 (such as embodied within a storage area network (SAN)).

[0026] The computer 200 is adapted to execute computer program modules for providing the functionality described herein. As used herein, the term “module” refers to computer program instruction and other logic for providing a specified functionality. A module can be implemented in hardware, firmware, and/or software. A module can include one or more processes, and/or be provided by only part of a process. A module is typically stored on the storage device 208, loaded into the memory 206, and executed by the processor 202.

[0027] The types of computer systems 200 used by the entities of FIG. 1 can vary depending upon the embodiment and the processing power used by the entity. For example, a user device 102 may be a mobile phone with limited processing power and a small display 218. The identification system 104, in contrast, may comprise multiple blade servers working together to provide the functionality described herein.

[0028] FIG. 3 is a block diagram illustrating modules operating within the identification

system 104 according to one embodiment. The identification system 104 includes a motion module 302, an area module 304, a fingerprint module 306, a search module 308, and a results module 310. The identification system 104 additionally includes a fingerprint storage 312 and a metadata storage 314. Those of skill in the art will recognize that other embodiments can have different and/or other modules than the ones described here, and that the functionalities can be distributed among the modules in a different manner.

[0029] The motion module 302 removes angular motion from video streams. When the identification system 104 receives a request for a video captured in a video stream to be identified, the motion module 302 analyzes the video frames of the stream for changes in camera angle/angular motion (vertical, horizontal and/or circular motion). Analysis of the video frames includes the motion module 302 comparing each frame of the stream with one or more frames immediately preceding it in the stream's sequence of frames. The motion module 302 determines whether vertical, horizontal, and/or circular motion occurred in the compared frame with respect to the one or more preceding frames. If vertical, horizontal, and/or circulation motion components are identified in the compared frame, the motion module 302 performs the necessary vertical, horizontal, and/or circular translation on the compared frame to remove the angular motion. Based on the translation of video frames that include angular motion, each frame of the stream appears to have been recorded by a stationary camera.

[0030] The area module 304 identifies display areas captured in video streams. After the motion module 302 removes angular motion from a stream capturing a video, the area module 304 analyzes the stream to identify a display area captured in the stream in which the video is displayed. The display area corresponds to the physical display in which playback of the video was displayed during the creation of the stream. For example, the display area may correspond to the display/screen of a television or a monitor. The display area needs to be identified so that it can be separated from the other portions of the stream and so that a fingerprint can be created that properly represents the video and not other content captured in the stream.

[0031] In one embodiment, the area module 304 identifies the top, bottom, left and right borders of the display area. To identify the top and bottom borders, the area module 304 analyzes each frame of the video stream from the top to the bottom (and/or bottom to the top) and identifies edges. These edges are referred to as top/bottom candidate edges.

[0032] For each top/bottom candidate edge, the area module 304 classifies the

candidate edge as interesting or boring based on the variety of brightness in the pixels of the edge. An interesting edge will have a variety of brightness within the edge, where a boring edge will not. To classify a top/bottom candidate edge as interesting or boring, the area module 304 determines the brightness level of each of the edge's pixels. Based on the brightness levels of the pixels, the area module 304 calculates a median brightness value for pixels of the edge. The area module 304 determines the number of pixels in the edge whose brightness level is within a brightness threshold (e.g., within 5 values) of the median brightness and the number edge pixels whose brightness level is not within the brightness threshold of the median.

[0033] In one embodiment, the area module 304 classifies the edge as boring if the number of pixels whose brightness level is within the brightness threshold of the median is greater than number of pixels whose brightness level is not within the threshold of the median. Otherwise the area module 304 classifies the edge as interesting. In another embodiment, the area module 304 classifies the edge as boring if number of pixels whose brightness level is not within the brightness threshold of the median is greater than a certain number. Otherwise the area module 304 classifies the edge as interesting.

[0034] For each top/bottom candidate edge, the area module 304 compares the interesting/boring classification given to the same edge in each frame to merge the classifications. If the top/bottom candidate edge is given the same classification in each frame, the area module 304 assigns the same classification to the edge. For example, if in each frame the edge is given the classification of boring, the area module 304 assigns the boring classification to the edge. However, if the classification given to the top/bottom candidate edge varies in the different frames, the area module 304 selects one of the classifications. In one embodiment, the area module 304 assigns to the edge, the classification given to the edge in majority of the frames. For example, if in majority of the frames the edge was classified as interesting, the edge is assigned an interesting classification. In another embodiment, if the classification given to the edge varies, the area module 304 assigns a default classification (e.g., a boring classification).

[0035] In another embodiment, instead of identifying top/bottom candidate edge in each frame and classifying each edge in each frame, the area module 304 blends the frames of the video stream to generate a single blended frame. The area module 304 identifies top/bottom candidate edges in the blended frame and classifies each edge as interesting or boring.

[0036] In addition to classifying each top/bottom candidate edge as interesting or

boring, the area module 304 also classifies each top/bottom candidate edge as tedious or exciting. Each top/bottom candidate edge is classified as tedious or exciting based on the variance of its pixels over time. Exciting edges have pixels that change over time, whereas tedious edges do not.

[0037] For each top/bottom candidate edge, the area module 304 determines the variance of each of the edge's pixels throughout the frames of the video stream. The area module 304 determines the number of pixels of the edge whose variance is less than a variance threshold (e.g., a value of 65) and the number of pixels whose variance is greater than the variance threshold. In one embodiment, the area module 304 classifies the top/bottom candidate edge as tedious if the number of pixels with variance less than the variance threshold is greater the number of pixels with variance greater than the threshold. Otherwise the area module 304 classifies the edge as exciting. In another embodiment, the area module 304 classifies the top/bottom candidate edge as tedious if the number of pixels with variance less than the variance threshold is greater than a certain number. Otherwise the area module 304 classifies the edge as exciting.

[0038] Based on the classifications of the top/bottom candidate edges, the area module identifies an approximate top border and an approximate bottom border. To identify the approximate top border, the area module 304 starts at the top of one the frames (e.g., Y value of zero of the first frame or a blended frame) and goes down the frame until it identifies a top/bottom candidate edge that has been classified as interesting and/or exciting. The area module 304 determines that the identified edge is the start of the display area because the edge has brightness variety (if classified as interesting) and/or varies over time (if classified as exciting). The area module 304 determines that the top/bottom candidate edge immediately above/before the identified edge on the Y-axis is the approximate top border.

[0039] The area module 304 performs the same process for the approximate bottom border but starts at the bottom of the frame and goes up until it identifies a top/bottom candidate edge classified as interesting and/or exciting. The area module 304 determines that the top/bottom candidate edge immediately below the identified edge on the Y-axis is the approximate bottom border. The area module 304 performs a Hough transform on the approximate top border and the approximate bottom border to identify the actual top border and bottom border of the display area.

[0040] To identify the left and right borders of the display area, the area module 304 rotates each frame of the video stream 90 degrees. The area module 304 repeats the process

used for identifying the top and bottom border to identify the left and right borders. In other words, the area module 304 identifies left/right candidate edges, classifies each left/right candidate edges as interesting or boring, classifies each left/right candidate edge as tedious or exciting, identifies an approximate left border and right border, and performs a Hough transform on the approximate borders to identify the left and right borders.

[0041] The area module 304 interconnects identified top, bottom, left, and right borders. The area enclosed by the interconnected borders is the display area in which the video is displayed in the captured video stream.

[0042] In another embodiment, instead of identifying the display area by identifying borders as described above, the area module 304 identifies the display area by analyzing motions within areas/regions and motion outside of these areas. In this embodiment, the area module 304 identifies multiple candidate areas in the frames of the video streams. For each candidate area, the area module 304 analyzes the amount of motion within candidate area throughout the frames of the stream and the amount of motion outside of the candidate area throughout the frames. The area module 304 selects a candidate area with motion within the area but little or no motion outside of the area as being the display area which the video is displayed in the stream.

[0043] In one embodiment, to select the candidate area, the area module 304 determines for each candidate area a candidate score which is a measure indicative of the amount of motion within the candidate area compared to the amount of motion outside the area. In one embodiment, the greater the amount of motion within the candidate area compared to outside the candidate area, the greater the candidate score. From the multiple candidate areas, the area module 304 selects the candidate area with the greatest candidate score as being the display area.

[0044] The fingerprint module 306 creates fingerprints for videos. When the area module 304 identifies a display area captured in a video stream, the fingerprint module 306 creates a fingerprint for the display area. To create the fingerprint, the fingerprint module 306 identifies each frame of the stream in which the display area is included. For each identified frame, the fingerprint module 306 crops the frame to remove from the frame content included outside of the display area. In one embodiment, if necessary, the fingerprint module 306 also performs perspective distortion on the display area if necessary.

[0045] The fingerprint module 306 identifies content included in the display area of each identified frame. The identified content includes one or more scenes of the video

captured in the stream. Based on the identified content, the fingerprint module 306 creates a fingerprint for the portion of the video captured in the video stream.

[0046] In one embodiment, to create the fingerprint, the fingerprint module 306 blurs together the identified content and finds maximally stable extremal regions. This results in the fingerprint module 306 generating a set of descriptors which it transforms into local quantized features. The local quantized features are turned into visterms, which are discrete representation of image characteristics and have weights. The weights of visterms are summed to produce the fingerprint.

[0047] The search module 308 searches for similar fingerprints in the fingerprint storage 312. The fingerprint storage 312 includes fingerprints for multiple videos. In one embodiment, the fingerprint storage 312 includes fingerprints of videos accessible from a video sharing system. Each fingerprint stored in the fingerprint storage 312 includes an identifier of a video to which the fingerprint corresponds.

[0048] For a fingerprint created by the fingerprint module 306 for a video, the search module 308 searches for fingerprints stored in the fingerprint storage 312 that are similar to the created fingerprint. The search module 308 identifies a certain number of fingerprints (e.g., one or three fingerprints) that are most similar to the created fingerprint. For each identified fingerprint, the search module 308 determines the video that corresponds to the fingerprint. The search module 308 provides identifiers of the videos that correspond to the identified fingerprints to the results module 310.

[0049] The results module 310 responds to requests from user devices 102 for identification of videos. When a request is received from a user device 102 for a video captured in a video stream to be identified, the results module 310 identifies any videos identified by search module 308 whose fingerprints are similar to the fingerprint created for the video captured in the stream.

[0050] The results module 310 determines for which of the identified videos to provide video information to the user device 102. In one embodiment, the results module 310 provides information about each video identified by the search module 308. In another embodiment, the results module 310 provides video information only for a select number of videos identified by the search module 308 (e.g., only for the identified video whose fingerprint is most similar to the fingerprint created by the fingerprint module 306).

[0051] For each identified video for which the results module 310 determines to provide video information, the results module 310 retrieves video information stored in the

metadata storage 314 about the video. The metadata storage 314 includes video information about each video for which one or more fingerprints are included in the fingerprint storage 312. The video information stored in the metadata storage 314 for a video in one embodiment may include one or more of the following: an identifier of the video (e.g., name of the video), the name of the video's publisher, a description of the video, a URL link for accessing the video from a system hosting the video, a number of times the video has been watched, a number of times people indicated they liked the video, and a number of times people indicated that they did not like the video.

[0052] The results module 310 transmits instructions to the user device 102 to display the retrieved video information to the user of the device 102. In one embodiment, if video information is retrieved for a single video, the video information is displayed as being information about the video whose playback was captured in the video stream. On the other hand, if video information is retrieved for multiple videos, the video information for each video is displayed as potentially being information about the video whose playback was captured in the video stream.

[0053] FIG. 4 is a flow chart illustrating operations 400 of the identification system 104 in identifying a video captured in a stream according to one embodiment. Those of skill in the art will recognize that other embodiments can perform the steps of FIG. 4 in different orders. Moreover, other embodiments can include different and/or additional steps than the ones described herein.

[0054] Assume for purposes of this example that a user uses a user device 102 to create a video stream that captures playback of a video. The identification system 104 receives 402 the video stream from the user device 102 along with a request for the video to be identified.

[0055] The identification system 104 removes 404 angular motion from video frames of the stream that include angular motion components. The identification system 104 identifies 406 a display area captured in the video frames, where the video is displayed within the display area in the frames. The identification system 104 creates 408 a fingerprint for the video based on content included in the display area throughout the frames of the stream.

[0056] From multiple stored fingerprints, the identification system 104 identifies 410 at least one stored fingerprint similar to the created fingerprint. The identification system 104 determines 412 a video to which the similar fingerprint corresponds. The identification system identifies 414 video information associated with the determined video, such as the name of the video and information for accessing the video from a system hosting the video.

The identification system 104 transmits 416 the video information to display to the user as video information about the video whose playback was captured in the stream.

[0057] The particular naming of the components, capitalization of terms, the attributes, data structures, or any other programming or structural aspect is not mandatory or significant, and the mechanisms that implement the embodiments described may have different names, formats, or protocols. Further, the systems may be implemented via a combination of hardware and software, as described, or entirely in hardware elements. Also, the particular division of functionality between the various system components described herein is merely exemplary, and not mandatory; functions performed by a single system component may instead be performed by multiple components, and functions performed by multiple components may instead performed by a single component.

[0058] Some portions of above description present features in terms of algorithms and symbolic representations of operations on information. These algorithmic descriptions and representations are the means used by those skilled in the data processing arts to most effectively convey the substance of their work to others skilled in the art. These operations, while described functionally or logically, are understood to be implemented by computer programs. Furthermore, it has also proven convenient at times, to refer to these arrangements of operations as modules or by functional names, without loss of generality.

[0059] Unless specifically stated otherwise as apparent from the above discussion, it is appreciated that throughout the description, discussions utilizing terms such as “processing” or “computing” or “calculating” or “determining” or “displaying” or the like, refer to the action and processes of a computer system, or similar electronic computing device, that manipulates and transforms data represented as physical (electronic) quantities within the computer system memories or registers or other such information storage, transmission or display devices.

[0060] Certain embodiments described herein include process steps and instructions described in the form of an algorithm. It should be noted that the process steps and instructions of the embodiments could be embodied in software, firmware or hardware, and when embodied in software, could be downloaded to reside on and be operated from different platforms used by real time network operating systems.

[0061] The embodiments described also relate to apparatuses for performing the operations herein. An apparatus may be specially constructed for the required purposes, or it may comprise a general-purpose computer selectively activated or reconfigured by a

computer program stored on a computer readable medium that can be accessed by the computer. Such a computer program may be stored in a non-transitory computer readable storage medium, such as, but is not limited to, any type of disk including floppy disks, optical disks, CD-ROMs, magnetic-optical disks, read-only memories (ROMs), random access memories (RAMs), EPROMs, EEPROMs, magnetic or optical cards, application specific integrated circuits (ASICs), or any type of media suitable for storing electronic instructions, and each coupled to a computer system bus. Furthermore, the computers referred to in the specification may include a single processor or may be architectures employing multiple processor designs for increased computing capability.

[0062] The algorithms and operations presented herein are not inherently related to any particular computer or other apparatus. Various general-purpose systems may also be used with programs in accordance with the teachings herein, or it may prove convenient to construct more specialized apparatus to perform the required method steps. The required structure for a variety of these systems will be apparent to those of skill in the, along with equivalent variations. In addition, the present embodiments are not described with reference to any particular programming language. It is appreciated that a variety of programming languages may be used to implement the teachings of the embodiments as described herein.

[0063] The embodiments are well suited for a wide variety of computer network systems over numerous topologies. Within this field, the configuration and management of large networks comprise storage devices and computers that are communicatively coupled to dissimilar computers and storage devices over a network, such as the Internet.

[0064] Finally, it should be noted that the language used in the specification has been principally selected for readability and instructional purposes, and may not have been selected to delineate or circumscribe the inventive subject matter. Accordingly, the disclosure of the embodiments is intended to be illustrative, but not limiting.

CLAIMS

1. A computer-implemented method comprising:
receiving, by a computer system from a user device, a request to identify a video;
receiving, by the computer system from the user device, a video stream capturing
playback of the video;
identifying, by the computer system, a display area captured in the video stream,
the display area corresponding to an object displaying the video as the
object is captured in the video stream;
creating, by the computer system, a fingerprint for the video based on content
included in the display area;
identifying, by the computer system, video information based on the created
fingerprint; and
transmitting, by the computer system to the user device, instructions to display the
video information.
2. The method of claim 1, wherein the object is a display displaying the video.
3. The method of claim 1, wherein identifying the display area comprises:
identifying a top border, a bottom border, a left border, and a right border of the
display area; and
interconnecting the identified borders.
4. The method of claim 1, wherein identifying the display area comprises:
identifying an area captured in the video stream with greater motion within the
area than outside the area throughout the stream.
5. The method of claim 1, wherein identifying the display area comprises:
identifying, a plurality of candidate areas in the video stream;
determining, for each of the plurality of candidate areas, a measure indicative of
an amount of motion within the candidate area compared to an amount of
motion outside of the candidate area throughout the plurality of frames;
and
selecting a candidate area from the plurality of candidates areas as the display area
based on the measure determined for the selected candidate area.

6. The method of claim 5, wherein the selected candidate area is selected based on having the greatest measure from the plurality of candidate areas.

7. The method of claim 1, wherein the display area is enclosed by four borders.

8. The method of claim 1, wherein identifying video information based on the fingerprint comprises:

identifying, from a plurality of stored fingerprints, at least one stored fingerprint similar to the created fingerprint; and

identifying video information associated with the at least one stored fingerprint, the identified video information displayed as being information about the video captured in the video stream.

9. The method of claim 1, wherein the video information comprises a video identifier and a link for accessing information associated with the video.

10. A computer program product having a non-transitory computer-readable storage medium having computer-executable instructions, the computer program product comprising:

a motion module configured to:

receive from a user device a request to identify a video; and
receive from the user device a video stream capturing playback of the video;

an area module configured to identify a display area captured in the video stream, the display area corresponding to an object displaying the video as the object is captured in the video stream;

a fingerprint module configured to create a fingerprint for the video based on content included in the display area; and

a results module configured to:

identify video information based on the created fingerprint; and
transmit to the user device instructions to display the video information.

11. The computer program product of claim 10, wherein the object is a display displaying the video.

12. The computer program product of claim 10, wherein the area module is further configured to:

identify a top border, a bottom border, a left border, and a right border of the display area; and

interconnect the identified borders.

13. The computer program product of claim 10, wherein the area module is further configured to:

identify an area captured in the video stream with greater motion within the area than outside the area throughout the stream.

14. The computer program product of claim 10, wherein the area module is further configured to:

identifying, a plurality of candidate areas in the video stream;

determining, for each of the plurality of candidate areas, a measure indicative of an amount of motion within the candidate area compared to an amount of motion outside of the candidate area throughout the plurality of frames; and

selecting a candidate area from the plurality of candidates areas as the display area based on the measure determined for the selected candidate area.

15. The computer program product of claim 14, wherein the selected candidate area is selected based on having the greatest measure from the plurality of candidate areas.

16. The computer program product of claim 10, wherein the display area is enclosed by four borders.

17. The computer program product of claim 10, further comprising:

a search module configured to identify, from a plurality of stored fingerprints, at least one stored fingerprint similar to the created fingerprint; and

the results module further configured to identify video information associated with the at least one stored fingerprint, the identified video information displayed as being information about the video captured in the video stream.

18. The computer program product of claim 10, wherein the video information comprises a video identifier and a link for accessing information associated with the video.

19. A system comprising:

one or more computer processors; and

one or more computer-readable storage mediums storing computer program modules adapted to execute on the one or more computer processors, the computer program modules comprising:

a motion module configured to:

- receive from a user device a request to identify a video; and
 - receive from the user device a video stream capturing playback of the video;
 - an area module configured to identify a display area captured in the video stream, the display area corresponding to an object displaying the video as the object is captured in the video stream;
 - a fingerprint module configured to create a fingerprint for the video based on content included in the display area; and
 - a results module configured to:
 - identify video information based on the created fingerprint; and
 - transmit to the user device instructions to display the video information.
20. The system of claim 19, wherein the area module is further configured to:
identify an area captured in the video stream with greater motion within the area than outside the area throughout the stream.

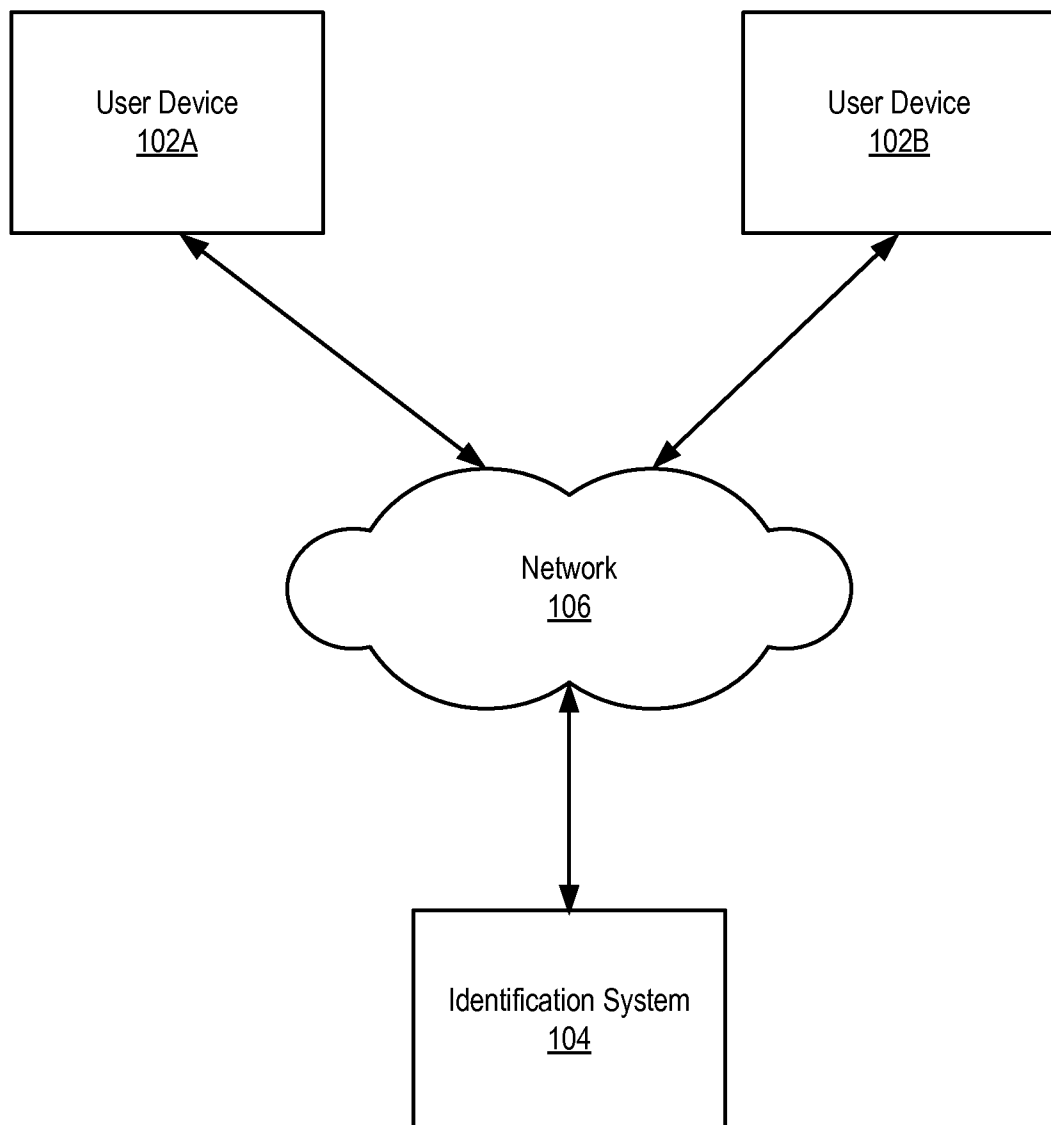
100

FIG. 1

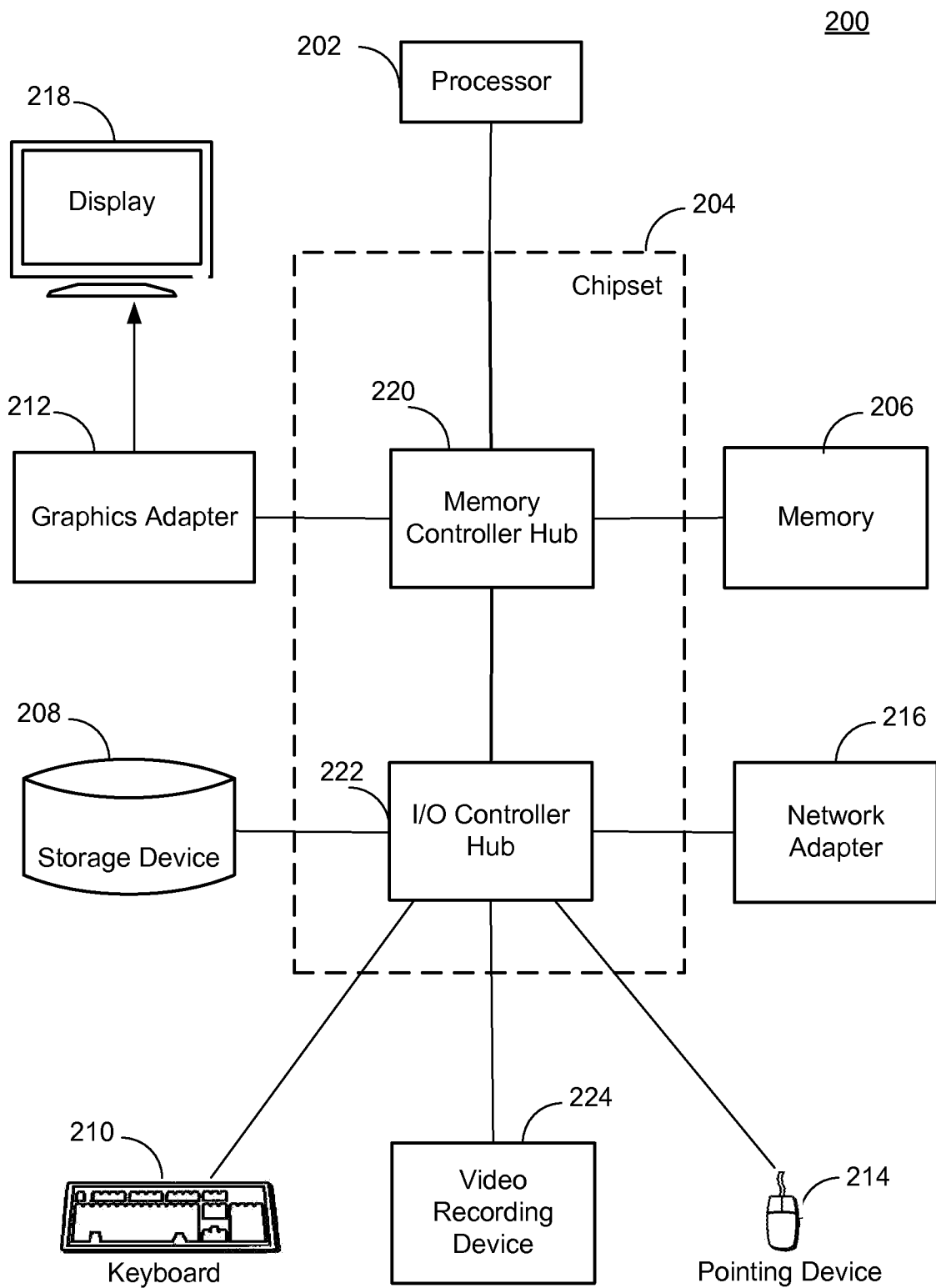


FIG. 2

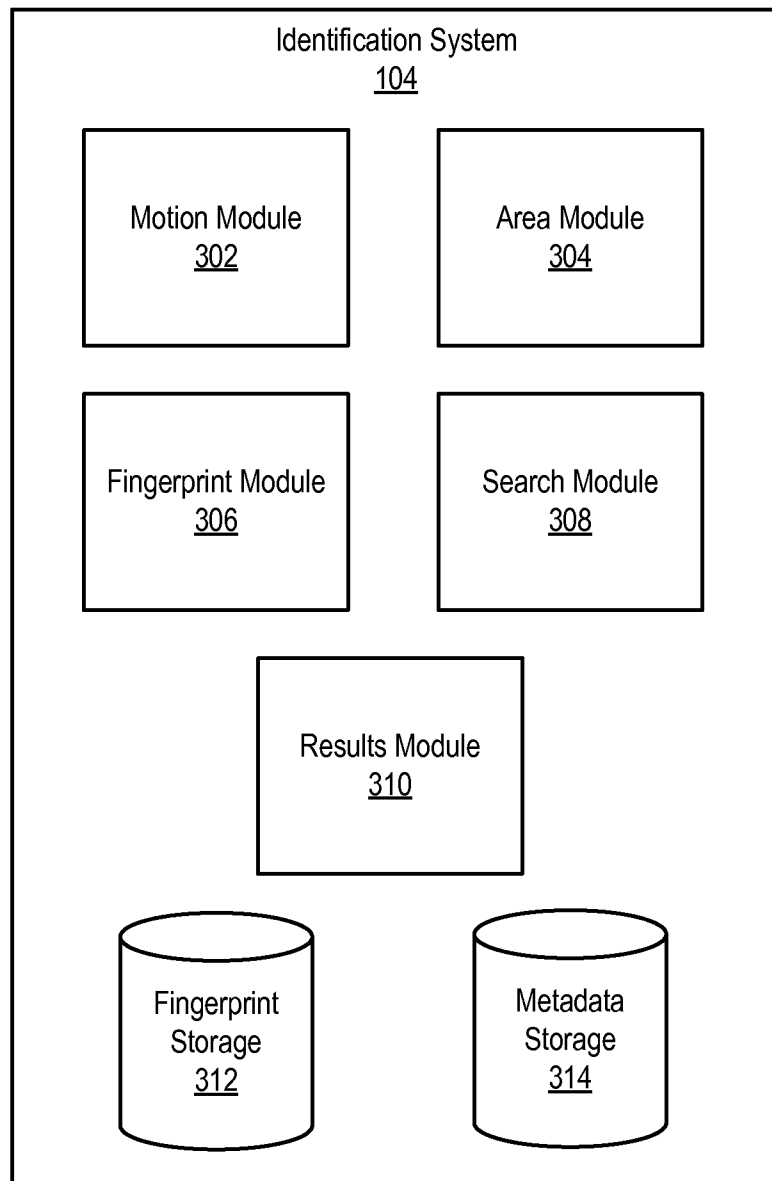


FIG. 3

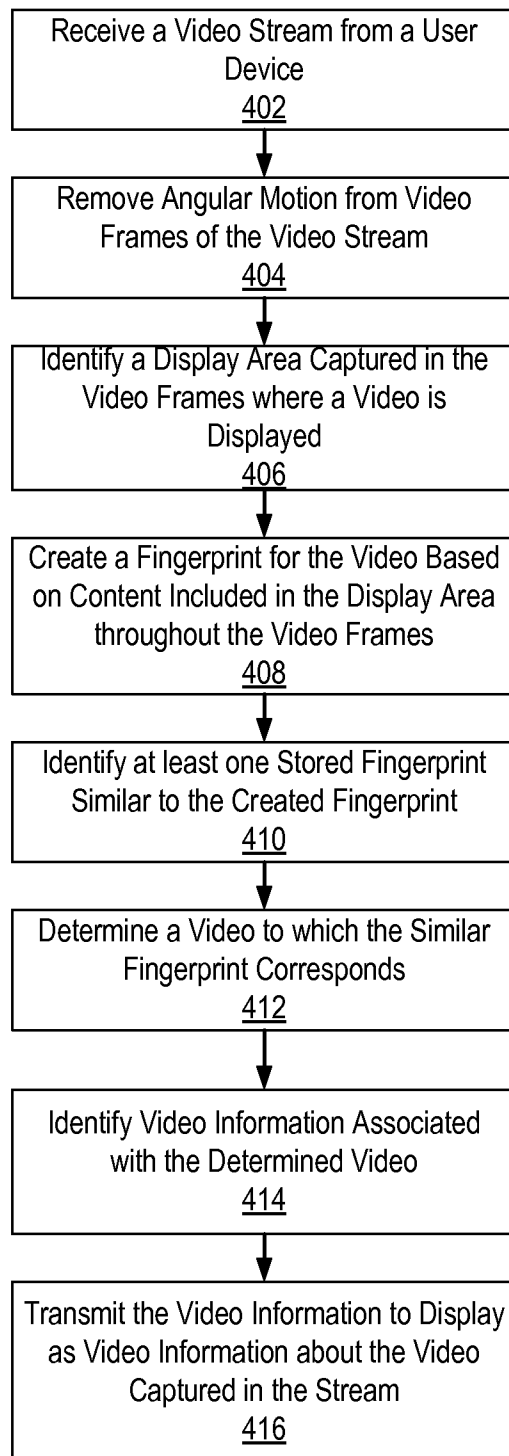
400

FIG. 4

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2015/044663

A. CLASSIFICATION OF SUBJECT MATTER
INV. H04N21/234 H04N21/4223 H04N21/278 G06K9/00 G06F17/30
ADD.

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 G06K G06F H04H

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)

EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2012/033876 A1 (MOMEYER BRIAN [US] ET AL) 9 February 2012 (2012-02-09) figures 1-4, 6, 7 paragraph [0032] - paragraph [0055] paragraph [0063] - paragraph [0070] -----	1-20
X	US 2011/311095 A1 (ARCHER DONALD GENE [US]) 22 December 2011 (2011-12-22)	1-3, 7-12, 16-19
Y	figures 1, 5, 8 paragraph [0016] - paragraph [0020] paragraph [0034] - paragraph [0055] paragraph [0065] - paragraph [0066] ----- -/-	4-6, 13-15,20



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

25 September 2015

Date of mailing of the international search report

06/10/2015

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Döttling, Martin

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2015/044663

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2014/236988 A1 (HARRON WILSON [US] ET AL) 21 August 2014 (2014-08-21)	1-3, 7-12, 16-19
Y	figures 1-4 paragraph [0025] - paragraph [0040] paragraph [0046] - paragraph [0048] paragraph [0053] - paragraph [0060] paragraph [0072] - paragraph [0082] -----	4-6, 13-15,20
Y	US 2014/044355 A1 (HARITA OGLU ISMAIL [US]) 13 February 2014 (2014-02-13) figure 7 paragraph [0020] paragraph [0089] - paragraph [0091] -----	4-6, 13-15,20
Y	US 2010/049711 A1 (SINGH GAJINDER [IN] ET AL) 25 February 2010 (2010-02-25) figures 2, 4 paragraph [0008] - paragraph [0015] paragraph [0036] - paragraph [0038] -----	1-20
Y	EP 2 391 122 A1 (NEC CORP [JP]) 30 November 2011 (2011-11-30) paragraphs [0022] - [0023] paragraph [0206] -----	1-20

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2015/044663

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2012033876 A1	09-02-2012	CN 103052961 A	17-04-2013
		EP 2601616 A2	12-06-2013
		JP 5666703 B2	12-02-2015
		JP 2013535752 A	12-09-2013
		KR 20130060280 A	07-06-2013
		US 2012033876 A1	09-02-2012
		WO 2012019163 A2	09-02-2012

US 2011311095 A1	22-12-2011	NONE	

US 2014236988 A1	21-08-2014	NONE	

US 2014044355 A1	13-02-2014	EP 2321964 A2	18-05-2011
		US 2011122255 A1	26-05-2011
		US 2014044355 A1	13-02-2014
		WO 2010011991 A2	28-01-2010

US 2010049711 A1	25-02-2010	US 2010049711 A1	25-02-2010
		US 2013259361 A1	03-10-2013

EP 2391122 A1	30-11-2011	CN 102292979 A	21-12-2011
		CN 104537127 A	22-04-2015
		EP 2391122 A1	30-11-2011
		EP 2434416 A2	28-03-2012
		JP 4883226 B2	22-02-2012
		JP 2012084164 A	26-04-2012
		KR 20110106414 A	28-09-2011
		US 2011274355 A1	10-11-2011
		WO 2010084737 A1	29-07-2010
