



(51) International Patent Classification:  
*H04N 7/16* (2011.01) *H04N 7/26* (2006.01)

(21) International Application Number:  
PCT/US2012/030396

(22) International Filing Date:  
23 March 2012 (23.03.2012)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
13/080,494 5 April 2011 (05.04.2011) US

(71) Applicant (for all designated States except US): **MICROSOFT CORPORATION** [US/US]; One Microsoft Way, Redmond, Washington 98052-6399 (US).

(72) Inventors: **HENDRICKSON, Seth**; c/o Microsoft Corporation, LCA - International Patents, One Microsoft Way, Redmond, Washington 98052-6399 (US). **DONG, George Randell**; c/o Microsoft Corporation, LCA - International Patents, One Microsoft Way, Redmond, Washington 98052-6399 (US). **MANDAL, Saswata**; c/o Microsoft Corporation, LCA - International Patents, One Microsoft Way, Redmond, Washington 98052-6399 (US). **ZITNICK III, Charles Lawrence**; c/o Microsoft Corporation, LCA - International Patents, One Microsoft Way, Redmond, Washington 98052-6399 (US).

[Continued on next page]

(54) Title: VIDEO SIGNATURE

(57) Abstract: Video signatures are generated when a small change in video information between consecutive frames N-1 and N is followed by a large change in video information between consecutive frames N and N+1. Information from frames N and/or N+1 is used to form a video signature.

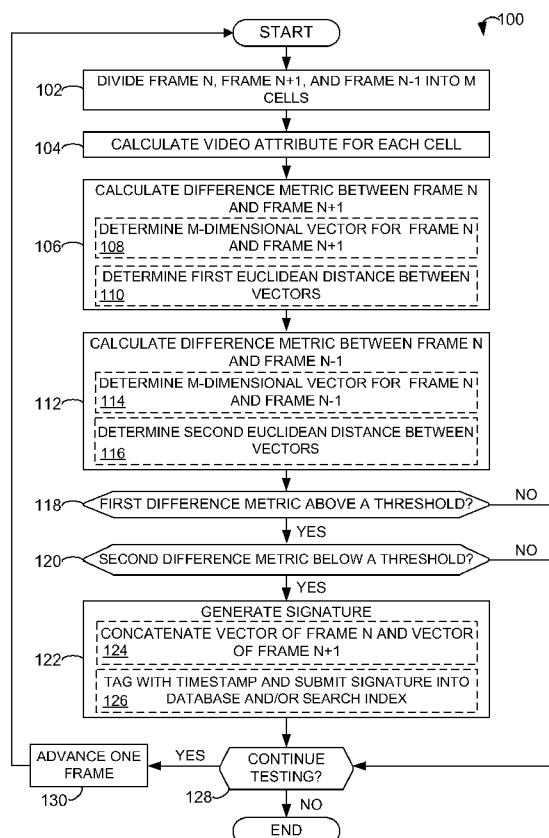


FIG.1



**(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, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

**(84) Designated States** (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD,

RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Declarations under Rule 4.17:**

- as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))
- as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))

**Published:**

- without international search report and to be republished upon receipt of that report (Rule 48.2(g))

## VIDEO SIGNATURE

### BACKGROUND

[0001] Identification information, such as metadata, is often not included with digital video content. Often this results in a viewer being unable to identify the title, year of production, genre of content, or any other information he or she may desire for a satisfactory viewing experience.

### SUMMARY

[0002] This Summary is provided to introduce a selection of concepts in a simplified form that are further described below in the Detailed Description. This Summary is not intended to identify key features or essential features of the claimed subject matter, nor is it intended to be used to limit the scope of the claimed subject matter. Furthermore, the claimed subject matter is not limited to implementations that solve any or all disadvantages noted in any part of this disclosure.

15 A method for finding a video signature is provided. The method comprises dividing each of frame N, frame N+1, and frame N-1 of the video into a plurality of cells and calculating a video attribute for each cell of each frame. A difference metric is calculated between frames N and N+1 and frames N and N-1 based on the video attributes of each frame, and the difference metrics are compared to a threshold. If the difference metric of frames N and N+1 exceeds a first threshold and a difference metric between frames N and N-1 does not exceed a second threshold, information from frame N and frame N+1 is used as a video signature.

### BRIEF DESCRIPTION OF THE DRAWINGS

25 [0003] FIG. 1 is a flow diagram depicting a method for finding a signature of a video stream according to one embodiment of the present disclosure.

[0004] FIG. 2 schematically shows frames from a video and associated luminance values according to one embodiment disclosed herein.

30 [0005] FIG. 3 is a flow diagram depicting an example embodiment for entering reference signatures into a search index database.

[0006] FIG. 4 is a flow diagram depicting an example embodiment for entering a client signature into a search index.

[0007] FIG. 5 schematically shows a non-limiting computing device for executing the embodiments disclosed herein.

## 5 DETAILED DESCRIPTION

[0008] The playback of media content, such as digital video, that lacks external identifying information may present problems, as the title of the video, the date it was produced, and other information may be needed by a viewer for a satisfactory viewing experience. To locate identifying information  
10 for an unknown video, video signatures may be generated by identifying shot changes within the video. Shot changes may be defined by a large change in the image information between two consecutive frames N and N+1 of the video, along with a small change in the image information between frame N and the preceding frame N-1. Information from the shot change can then be  
15 used to generate a video signature. The video signature may be easily reproducible on different versions of the same video.

[0009] Such a video signature may be compared to a video signature database that includes video signatures for a library of reference videos, as well as identifying information (e.g., metadata) for each such reference video.  
20 If a match is found, the external identifying information from the matching reference video can be applied to the unknown video. In some embodiments, a Euclidean distance between the tested signature and the reference signature may be used to determine whether a match exists. In some embodiments, such a Euclidean distance must be zero, and in other embodiments a small  
25 difference may be allowed.

[0010] FIG. 1 shows a flow diagram depicting a method 100 for finding a video signature according to one embodiment of this disclosure. Method 100 may be executed by a computing device capable of video playback and/or video streaming. Example computing devices include but are not limited to desktop  
30 computers, laptop computers, mobile computing devices, and TV set-top boxes. The computing device may further be connected to other computing devices over a network for receiving video streams and/or accessing a video signature

database, which will be described in more detail below. Further detail about the computing device will be described with reference to FIG. 5.

[0011] At 102, method 100 includes dividing frames within the video into a plurality of cells. Each frame of a series of frame N, frame N+1, and frame N-1 are divided into M cells. In one example, M cells may include 16 equal sized cells arranged in a four by four (4x4) grid. In another example, M cells may include 25 equal sized cells arranged in a five-by-five grid. Any number of cells in any arrangement is within the scope of this disclosure.

[0012] A video attribute for each cell is calculated at 104. In some embodiments, the video attribute is an average luminance for that cell, calculated by averaging the luminance values for each pixel within the cell. In some embodiments, the video attribute may be an average luminance for that cell normalized with respect to an average luminance for all cells of that frame. Other suitable attributes could be used, such as rank order of the luminance of each cell, an average color value, etc.

[0013] At 106, a first difference metric is calculated between frame N and frame N+1. As an example, the difference metric may be a Euclidean distance between a vector calculated for frame N and a vector calculated for frame N+1. To calculate this difference metric, an M-dimensional vector for each of frame N and frame N+1 is determined at 108. The M-dimensional vector for each frame includes, for each different dimension, a video attribute of a different cell from that frame.

[0014] For example, FIG. 2 shows a frame N of a video divided into a 4x4 grid including 16 cells (i.e., cell N1, N2, ... N16). Furthermore, table 204 of FIG. 2 shows a luminance value for each cell of frame N. For each cell, the average luminance is calculated from the luminance value for each pixel of that cell. In this example, luminance is represented by a scale from 0 to 1, where 0 is no luminance (i.e. the cell is all black and no light is emitted) and 1 is the maximum luminance (i.e. the cell is all white and the maximum amount of light the pixels can emit is being emitted). For example, cell N1 has a luminance value of 0.2. The M-dimensional vector may include a 16-dimensional vector based on the luminance for each of the 16 cells of the

divided frame. For example, frame N would have a 16-dimensional vector comprising:

$$\vec{v}_N = (0.2, 0.98, 0.95, 0.95, 0.5, 0.9, 0.9, 0.95, 1, 0.99, 0.99, 1, 1, 1, 1, 1)$$

while frame N+1, based on the table of luminance values 206, would have a 16-dimensional vector comprising:

$$\vec{v}_{N+1} = (0.98, 0.97, 0.97, 0.98, 0.97, 1, 1, 0.97, 0.95, 0.5, 0.5, 0.95, 0.9, 0.75, 0.75, 0.9).$$

Turning back to FIG. 1, at 110, method 100 includes calculating a Euclidean distance between the vectors. For example, in FIG. 2, the Euclidean distance between frame N and frame N+1 of FIG. 2 may be calculated using the following equation:

$$d(p, q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 \dots (p_n - q_n)^2}$$

where p is the vector for frame N, q is the vector for frame N+1, and n=16 as the vectors comprise 16 dimensions. Therefore, the Euclidean distance between frame N and frame N+1 of FIG. 2 would be approximately 1.21.

[0015] Referring back to FIG. 1, a second difference metric is calculated between frame N and frame N-1 at 112. Similar to the difference metric calculated at 106, an M-dimensional vector is determined for each frame N and frame N-1 at 114, and these vectors are used to determine a second Euclidean distance at 116.

[0016] At 118, the first difference metric calculated between frames N and N+1 is compared to a first threshold. If the first difference metric is above the first threshold, the method proceeds to 120. If it is not above the threshold, the method advances to 128. The second difference metric is compared to a second threshold at 120. If the second difference metric is below the second threshold, the method proceeds to 122, and if it is not below the threshold, the method advances to 128. The first and second thresholds for comparing the first and second difference metrics may be similar. Alternatively, the two thresholds may be different. The thresholds may be determined empirically or by any suitable mechanism.

[0017] Referring to the example frames depicted in FIG. 2, for frame N-1, a 16-dimensional vector based on the luminance values in table 202 would be the same as the vector for frame N, as the frames depict the same image.

Therefore, the Euclidean distance, calculated using the same equation as described above, would result in a distance of 0 between frames N and N-1.

[0018] In order to determine if the frames constitute a shot change, the calculated Euclidean distances can each be compared to a threshold. In this example, the first and second thresholds are equal, and empirically determined to be equal to one. Thus, the distance from frame N to frame N+1,  $\Delta_{N \rightarrow N+1} = 1.21$ , which is greater than the first threshold of 1. The distance from frame N to frame N-1,  $\Delta_{N \rightarrow N-1} = 0$ , which is less than the second threshold, which is also 1 in this case. As the first distance is greater than the threshold and the second is not, a shot change is identified.

[0019] At 122 in FIG. 1, information from one or both of frame N and frame N+1 is included as a video signature. The vector of frame N may be concatenated with the vector of frame N+1 at 124 to generate the signature, in this embodiment a 32-dimensional vector based on the luminance values for frame N and frame N+1. At 126, the signature is associated with a timestamp indicating the time relative to the start of the video at which the shot change occurs. The signature may then be submitted to a database of predetermined signatures, and/or the signature may be entered into a search index to search for a matching reference signature. For example, if the video is a known video that contains identifying information, the signature may be submitted to a reference video signature database along with the identifying information for use at a later time, such as when searching the database for a match to an unknown video. Alternatively, if the video is an unknown video that does not contain identifying information, the signature generated may be entered into a search index to search for a matching signature contained within a known reference video from the reference video signature database.

[0020] Method 100 may proceed to iteratively test subsequent frames if it is determined at 128 that the testing should continue to identify other signatures within the video. Virtually any criteria may be used to determine if additional testing should continue. For example, such testing may continue unless the end of the video has been reached or a user indicates that testing should be terminated. If the testing is to continue, the method advances one

frame at 130, and loops back to the beginning of method 100. Additionally, if either of the first or second difference metrics does not meet its respective condition relative to the thresholds at 118 and 120, one frame may be advanced and the method may iteratively test subsequent frames. It should  
5 be understood that some steps for subsequent testing may not need to be repeated from previous testing. For example, if one of the tested frames has already been divided and an M-dimensional matrix has been found for that frame, such dividing and calculating need not be repeated.

[0021] Alternatively, if it is determined the testing should not continue  
10 (e.g. if all signatures in the video have been identified) the method ends.

[0022] It should be understood that the method 100 need not be carried out in the order described. For example, the difference metric between frame N and frame N-1 may be calculated before, after, or concurrently to the calculation of the difference metric between frame N and frame N+1.  
15 Similarly, comparing the first difference metric to the threshold may occur before, after, or concurrently to the comparison of the second difference metric to the threshold.

[0023] By using average luminance values of a grid of even sized cells, shot changes may be identified even when manipulations have been made to  
20 the video. For example, it is common in videos distributed over the internet for the contrast, brightness, etc. to be altered from the original video source. By utilizing an average luminance for each cell normalized to the average luminance for the entire frame, differences in relative brightness may be accounted for. Furthermore, the identification of shot changes is not  
25 dependent on the timing of the video, so that videos that have been sped up or slowed down may also be identified. Finally, clips of videos may also be identified, so long as the clip contains at least one shot change.

[0024] The process of identifying shot changes and creating signatures can be performed on video content that contains identifying information as  
30 well as video content that does not contain identifying information. Video signatures from a known video can be associated with that video in a reference database. One example method for searching such a database includes



locality sensitive hashing. In this method, each reference signature is given a code, termed a hash code, and each hash code is distributed into buckets of a hash table. The unknown video signature, or client signature, is also given a hash code. The bucket the client signature is distributed to reasonably  
 5 contains only those reference signatures similar to the client signature. From these, a difference metric, such as a Euclidean distance, may be calculated between the client signature and the references signatures.

[0025] An example method 300 of generating hash codes for reference video signatures is described with reference to FIG. 3. At 302, method 300  
 10 includes generating a set of  $n$  random vectors ( $v_1, v_2 \dots v_n$ ), the components of which should be independent and identically distributed normal random variables. The  $n$  random vectors may be generated once, and the set may be used for determining a hash code at 304 for every signature generated. The following nonlimiting example will be described with reference to a 32-  
 15 dimensional vector corresponding to frames divided into 4x4 grids.

[0026] A hash  $H_i$  is determined at 306 for each random vector. Each  $H_i$  may be calculated using the following equation:

$$H_i = \frac{[S \cdot v_i]}{B}$$

where  $S$  is the 32-dimensional reference signature,  $v_i$  is the 32-dimensional random vector, and  $B$  is the bucket size, which may be determined  
 20 experimentally. A large  $B$  may produce more reliable search results, however it increases the size of the hash bucket and thus the length of time required to perform a search within that bucket due to the increased number of signatures within that bucket. At 308, the hash may then be converted to binary to get a hash code.  $H_i$  is determined for  $i = 1, 2 \dots n$ , and then all are  
 25 converted to binary and concatenated to get a super hash code  $H$  at 310. Method 300 includes, at 312, indexing each super hash code into a bucket of a reference hash table along with tags, such as a timestamp from the video where the signature occurs, title of video, etc. If more reference signatures are to be hashed at 314, the process repeats itself.

[0027] FIG. 4 depicts a flow diagram showing an embodiment of a method 400 for entering an unknown, or client, signature into a search index to find a matching reference signature. The method 400 includes determining a query hash code for the client signature at 402. Similar to the method described in reference to FIG. 3, a hash is determined at 406 based on the client signature, random vector (retrieved at 404 from the set generated at 302), and bucket size, for each random vector of the set. Each hash code is converted to binary at 408, and all binary hash codes are concatenated at 410 to produce a query super hash code for the client signature, H'. The query super hash code H' is assigned to a bucket of the reference hash table, and all reference hashes stored in that bucket are returned at 412. A difference metric, such as a Euclidean distance, can be calculated at 414 between the client signature and each reference signature in the bucket. At 416, all reference signatures that have Euclidean distances below a threshold may be returned as a match. If no matches exist at 418, the method may proceed to 424. If more than one match exists at 420, the signature may be deemed to derive from common content, such as a commercial, or be a low quality signature, and thus the method may advance to 424. If it is determined at 418 and 420 that only one match exists, there is a high probability the client signature and reference signature are from the same video. In this case, the client video may be identified as the reference video at 422, and the information accompanying the reference video, such as title, date, etc., may be added to the client video. At 424, method 400 includes determining if it is necessary to continue searching signatures within the video for more matches. For example, it may be necessary to keep searching for matches to additional client signatures within the video if no matches could initially be found, or if multiple matches were found. If it is determined that searching should be continued, the method 400 advances a next signature in the video at 426 and loops back to the beginning of method 400.

[0028] Additionally, accuracy of the matching beyond the difference metric between the client signature and the reference signature may be increased. This can be accomplished by verifying that more than one client

signature in the video matches a reference signature from that video. Further, if multiple matching signatures are found between the known and unknown video, the accuracy can be further verified by correlating the difference between the timestamps from two signatures in the reference video  
5 vs. the difference between the timestamps of the same two signatures of the client video. As explained above with regard to FIG. 1, timestamps may be associated with the video signatures. As the client video content may be altered compared to the reference video, using an absolute time correlation may not be an accurate method of verification, as the client video may be  
10 shortened, or include extra video information. However, if the difference between two timestamps is used, an accurate correlation may be made. For example, a first video signature of a reference video may be associated with timestamp 1:35 and a second signature of that video may be associated with 1:50. In this example a difference between the two is 0:15. If the client video  
15 has a signature that matches the first reference signature with a timestamp of 2:15 and a second signature that matches the second reference signature with a timestamp of 2:30, there is a high likelihood that the matches are accurate, as there are two matching signatures, and the differences between the timestamps of the signatures are the same.

20 **[0029]** In some embodiments, the above described methods and processes may be tied to a computing system including one or more computers. In particular, the methods and processes described herein may be implemented as a computer application, computer service, computer API, computer library, and/or other computer program product.

25 **[0030]** FIG. 5 schematically shows a nonlimiting computing system 502 that may perform one or more of the above described methods and processes. Computing system 502 is shown in simplified form. It is to be understood that virtually any computer architecture may be used without departing from the scope of this disclosure. In different embodiments, computing system 502 may  
30 take the form of a mainframe computer, server computer, desktop computer, laptop computer, tablet computer, home entertainment computer, network

computing device, mobile computing device, mobile communication device, gaming device, etc.

[0031] Computing system 502 includes a logic subsystem 504 and a data-holding subsystem 506. Computing system 502 may optionally include a display subsystem 508, communication subsystem 510, and/or other components not shown in FIG. 5. Computing system 502 may also optionally include user input devices 512 such as keyboards, mice, game controllers, cameras, microphones, and/or touch screens, for example.

[0032] Computing system 502 may be configured to communicate with a remote computing device such as a server 514 over a network 516. Similar to computing system 502, server 514 includes a logic subsystem 518 and a data-holding subsystem 520. Server 514 may send video streams to computing system 502 for viewing by a user. The video streams, if they lack identifying information, may be searched for video signatures, as described with reference to FIGS. 1 and 2. Furthermore, server 514, or another server, may search reference videos for video signatures and also may utilize a reference signature database 522 for storing tagged reference signatures. Alternatively, a reference signature database 524 may be stored remotely from the server, and the server may access the database over the network.

[0033] Because of the distributed nature of a hash table, the duties of server 514 may be split up into a distributed network of servers. In this case, there would be a load balancer which communicates with computing system 502, and delegates signature matching to one of several servers depending on the hash value for the incoming signature match request. Each server would host a subset of the database corresponding to certain hash buckets from the hash table. There could optionally be overlap in which hash buckets each server holds, but each hash bucket will be held by some server, so that the entire hash table is represented throughout the distributed system.

[0034] Logic subsystem 504 may include one or more physical devices configured to execute one or more instructions. For example, the logic subsystem may be configured to execute one or more instructions that are part of one or more applications, services, programs, routines, libraries, objects,

components, data structures, or other logical constructs. Such instructions may be implemented to perform a task, implement a data type, transform the state of one or more devices, or otherwise arrive at a desired result.

[0035] The logic subsystem may include one or more processors that are  
5 configured to execute software instructions. Additionally or alternatively, the logic subsystem may include one or more hardware or firmware logic machines configured to execute hardware or firmware instructions. Processors of the logic subsystem may be single core or multicore, and the programs executed thereon may be configured for parallel or distributed processing. The  
10 logic subsystem may optionally include individual components that are distributed throughout two or more devices, which may be remotely located and/or configured for coordinated processing. One or more aspects of the logic subsystem may be virtualized and executed by remotely accessible networked computing devices configured in a cloud computing configuration.

15 [0036] Data-holding subsystem 506 may include one or more physical, non-transitory, devices configured to hold data and/or instructions executable by the logic subsystem to implement the herein described methods and processes. When such methods and processes are implemented, the state of data-holding subsystem 506 may be transformed (e.g., to hold different data).

20 [0037] Data-holding subsystem 506 may include removable media and/or built-in devices. Data-holding subsystem 506 may include optical memory devices (e.g., CD, DVD, HD-DVD, Blu-Ray Disc, etc.), semiconductor memory devices (e.g., RAM, EPROM, EEPROM, etc.) and/or magnetic memory devices (e.g., hard disk drive, floppy disk drive, tape drive, MRAM, etc.),  
25 among others. Data-holding subsystem 506 may include devices with one or more of the following characteristics: volatile, nonvolatile, dynamic, static, read/write, read-only, random access, sequential access, location addressable, file addressable, and content addressable. In some embodiments, logic subsystem 504 and data-holding subsystem 506 may be integrated into one or  
30 more common devices, such as an application specific integrated circuit or a system on a chip.

[0038] FIG. 5 also shows an aspect of the data-holding subsystem in the form of removable computer-readable storage media 526, which may be used to store and/or transfer data and/or instructions executable to implement the herein described methods and processes. Removable computer-readable storage media 526 may take the form of CDs, DVDs, HD-DVDs, Blu-Ray Discs, EEPROMs, and/or floppy disks, among others.

[0039] It is to be appreciated that data-holding subsystem 506 includes one or more physical, non-transitory devices. In contrast, in some embodiments aspects of the instructions described herein may be propagated in a transitory fashion by a pure signal (e.g., an electromagnetic signal, an optical signal, etc.) that is not held by a physical device for at least a finite duration. Furthermore, data and/or other forms of information pertaining to the present disclosure may be propagated by a pure signal.

[0040] The terms “module,” “program,” and “engine” may be used to describe an aspect of computing system 502 that is implemented to perform one or more particular functions. In some cases, such a module, program, or engine may be instantiated via logic subsystem 504 executing instructions held by data-holding subsystem 506. It is to be understood that different modules, programs, and/or engines may be instantiated from the same application, service, code block, object, library, routine, API, function, etc. Likewise, the same module, program, and/or engine may be instantiated by different applications, services, code blocks, objects, routines, APIs, functions, etc. The terms “module,” “program,” and “engine” are meant to encompass individual or groups of executable files, data files, libraries, drivers, scripts, database records, etc.

[0041] It is to be appreciated that a “service”, as used herein, may be an application program executable across multiple user sessions and available to one or more system components, programs, and/or other services. In some implementations, a service may run on a server responsive to a request from a client.

[0042] When included, display subsystem 508 may be used to present a visual representation of data held by data-holding subsystem 506. As the herein described methods and processes change the data held by the data-holding subsystem, and thus transform the state of the data-holding subsystem, the state of display subsystem 508 may likewise be transformed to visually represent changes in the underlying data. Display subsystem 508 may include one or more display devices utilizing virtually any type of technology. Such display devices may be combined with logic subsystem 504 and/or data-holding subsystem 506 in a shared enclosure, or such display devices may be peripheral display devices.

[0043] When included, communication subsystem 510 may be configured to communicatively couple computing system 502 with one or more other computing devices, such as server 514. Communication subsystem 510 may include wired and/or wireless communication devices compatible with one or more different communication protocols. As nonlimiting examples, the communication subsystem may be configured for communication via a wireless telephone network, a wireless local area network, a wired local area network, a wireless wide area network, a wired wide area network, etc. In some embodiments, the communication subsystem may allow computing system 502 to send and/or receive messages to and/or from other devices via a network such as the Internet.

[0044] It is to be understood that the configurations and/or approaches described herein are exemplary in nature, and that these specific embodiments or examples are not to be considered in a limiting sense, because numerous variations are possible. The specific routines or methods described herein may represent one or more of any number of processing strategies. As such, various acts illustrated may be performed in the sequence illustrated, in other sequences, in parallel, or in some cases omitted. Likewise, the order of the above-described processes may be changed.

30

[0045] The subject matter of the present disclosure includes all novel and nonobvious combinations and subcombinations of the various processes, systems and configurations, and other features, functions, acts, and/or properties disclosed herein, as well as any and all equivalents thereof.

5



## CLAIMS:

1. A method of finding a signature for a video, the method comprising:

dividing a frame N of the video into a plurality of cells;

calculating a video attribute of each cell of frame N;

dividing a frame N+1 into the plurality of cells;

calculating a video attribute of each cell of frame N+1;

calculating a difference metric between frame N and frame N+1 that includes, for each cell of the plurality of cells, a difference between the video attribute of frame N and the video attribute of frame N+1;

if the difference metric between frame N and frame N+1 exceeds a first threshold and a difference metric between frame N and frame N-1 does not exceed a second threshold, including information from one or both of the frame N and the frame N+1 as a video signature.

2. The method of claim 1, wherein the plurality of cells includes M cells, and where an M-dimensional vector for each frame includes the video attribute for each of the plurality of M cells.

3. The method of claim 2, wherein calculating the difference metric between frame N and frame N+1 includes calculating a Euclidean distance between the M-dimensional vector for frame N and the M-dimensional vector for frame N+1.

4. The method of claim 1, where the video attribute is an average luminance for that cell.

5. The method of claim 1, where the video attribute is an average luminance for that cell normalized with respect to an average luminance for all cells of that frame.

6. The method of claim 1, where the video signature includes each of the video attributes of frame N concatenated with each of the video attributes of frame N+1.

7. The method of claim 1, where dividing the frame N of the video into a plurality of cells includes dividing the frame N into a four by four grid of cells.

8. The method of claim 1, wherein frame N-1 is divided into M cells and an M-dimensional vector for frame N-1 includes the video attribute for each cell, and wherein calculating the difference metric between frame N and frame N-1 includes calculating a Euclidean distance between the M-dimensional vector for frame N and the M-dimensional vector for frame N-1.

9. The method of claim 1, comprising associating a timestamp with the video signature.

10. A data-holding subsystem holding instructions executable by a logic subsystem to:

iteratively test a series of three consecutive frames within a video for a shot change, each test comprising:

dividing a frame N of the video into a grid comprising M cells;

calculating an M-dimensional vector for frame N that includes, for each different dimension of the M-dimensional vector, an average luminance value of a different cell from frame N;

dividing a frame N+1 of the video into a grid comprising M cells;

calculating an M-dimensional vector for frame N+1 that includes, for each different dimension of the M-dimensional vector, an average luminance value of a different cell from frame N+1;

dividing a frame N-1 of the video into a grid comprising M cells;

calculating an M-dimensional vector for frame N-1 that includes, for each different dimension of the M-dimensional vector, an average luminance value of a different cell from frame N-1;

calculating a Euclidean distance between the M-dimensional vector for frame N and the M-dimensional vector for frame N+1;

if the Euclidean distance between frame N and frame N+1 exceeds a first threshold, calculating a Euclidean distance between the M-dimensional vector for frame N and the M-dimensional vector for frame N-1; and

if the Euclidean distance between frame N and frame N-1 does not exceed a second threshold, including the M-dimensional vector for frame N concatenated with the M-dimensional vector of frame N+1 as a video signature, and

iteratively test a next series of three consecutive frames within the video for a next shot change.

1/5

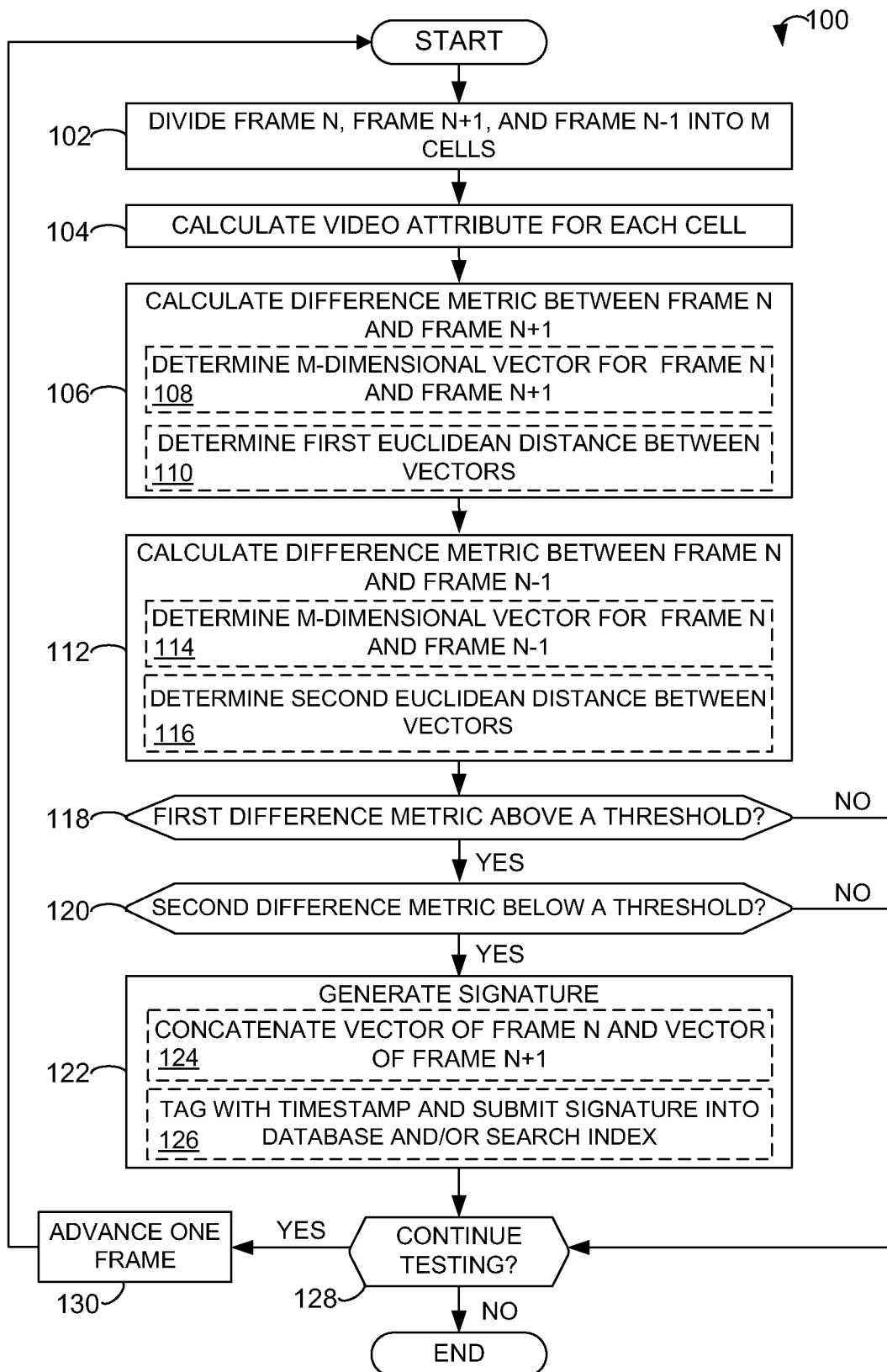


FIG.1

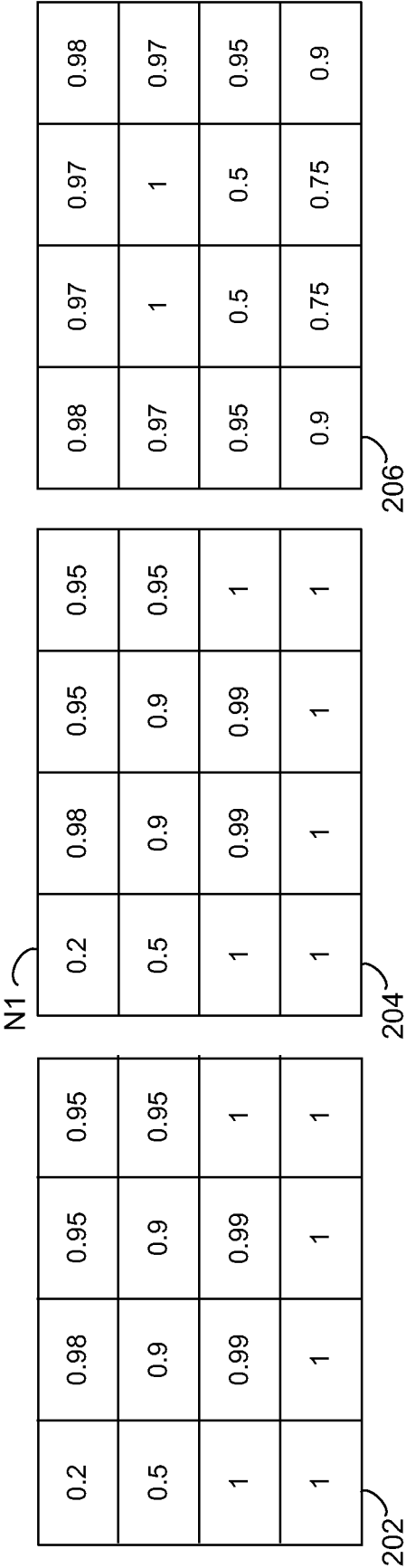
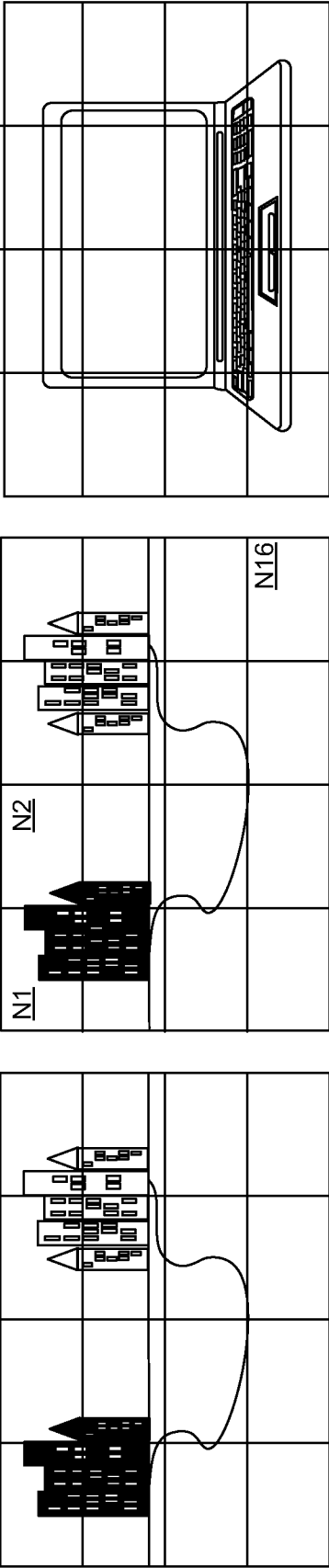
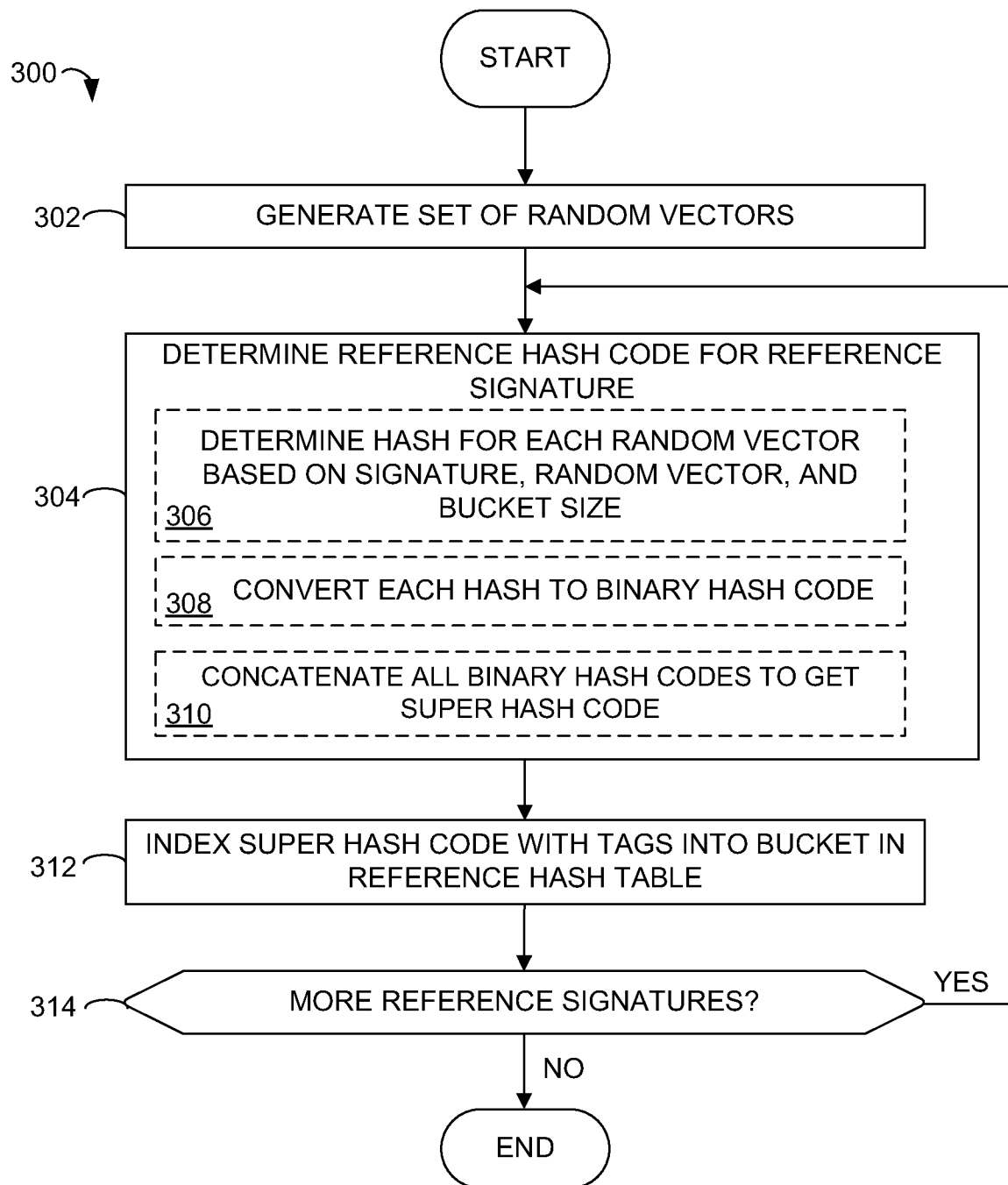


FIG. 2

3/5

**FIG. 3**

4/5

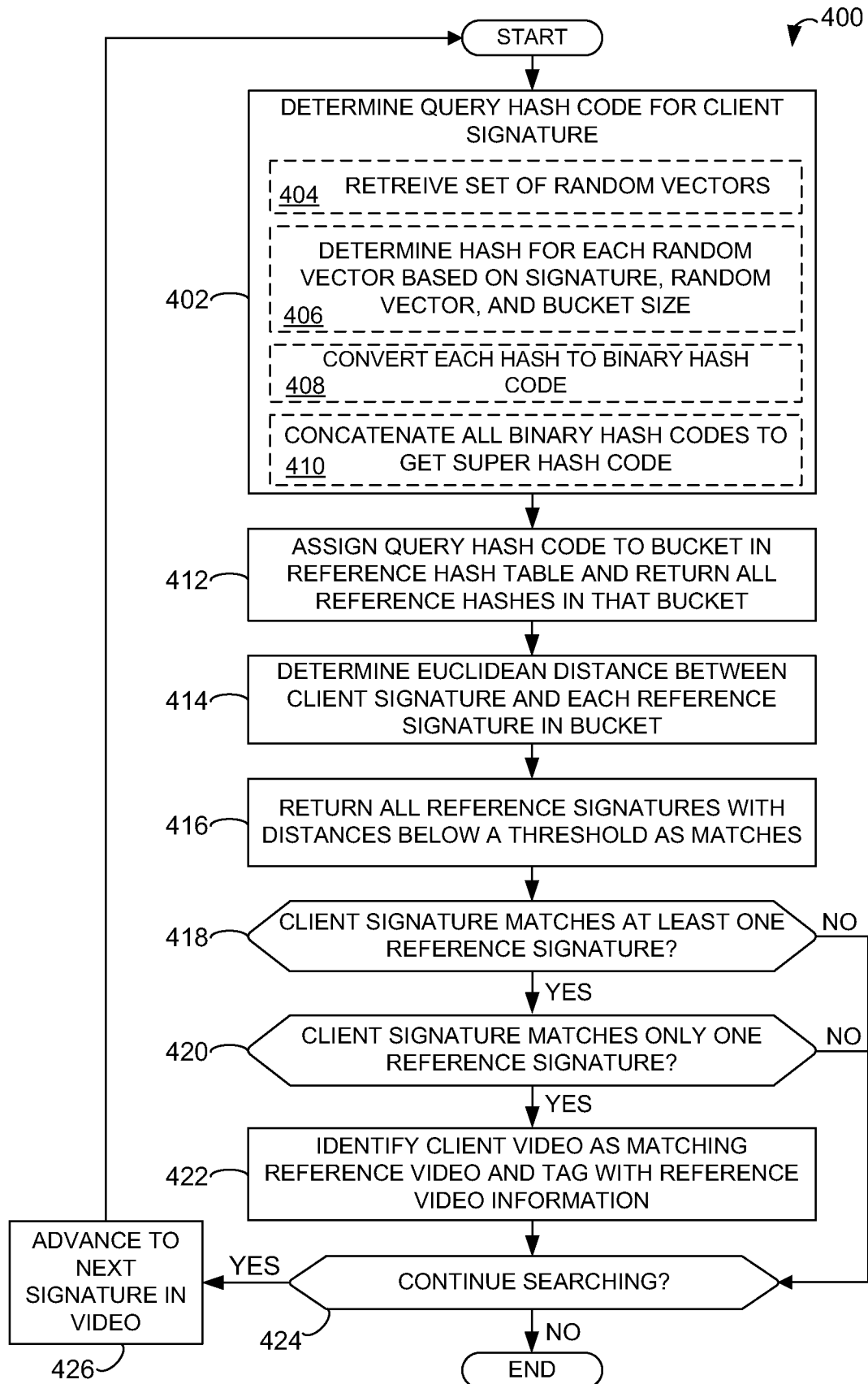
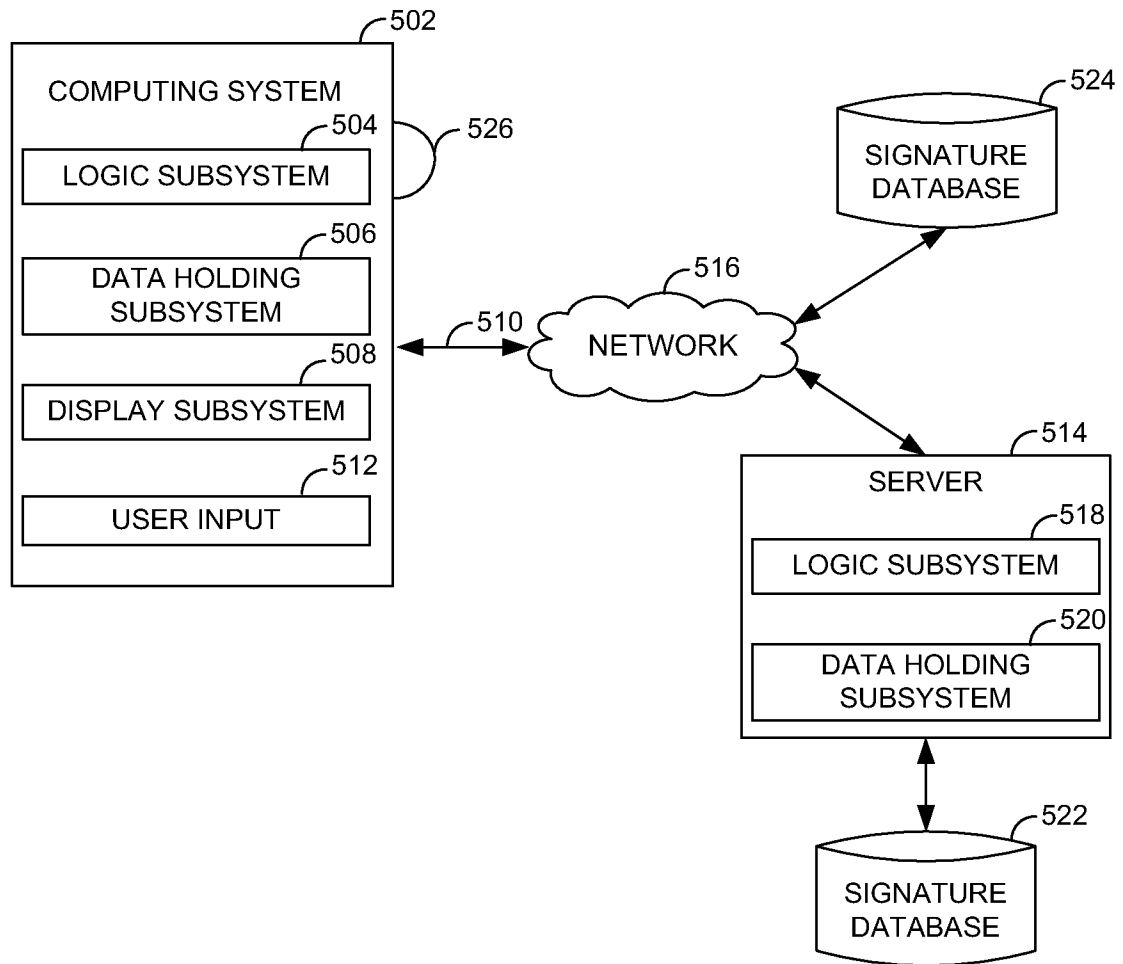


FIG. 4

5/5

**FIG. 5**