

19



NL Octrooi Centrum

11

2007887

12 C OCTROOI

21 Aanvraagnummer: **2007887**

51 Int.Cl.: **G06F 17/30** (2006.01) **G06F 3/048** (2006.01)

22 Aanvraag ingediend: **29.11.2011**

30 Voorrang:
09.05.2011 US 13/103996

43 Aanvraag gepubliceerd:
-

47 Octrooi verleend:
13.11.2012

45 Octrooischrift uitgegeven:
21.11.2012

73 Octrooihouder(s):
Google Inc. te MOUNTAIN VIEW, California, Verenigde Staten van Amerika (US).

72 Uitvinder(s):
**Shivakumar Rajaraman te Mountain View, California (US).
Vasilios Pantazopoulos te Mountain View, California (US).
John Skidgel te Mountain View, California (US).**

74 Gemachtigde:
mr. ir. J. van Breda c.s. te Amsterdam.

54 **Contextual video browsing.**

57 A method and system for contextual browsing of videos that are relevant to a current video. Browsing controls that include labels and previews are presented in a user interface. The labels represent sets of videos that are organized according to their relevancy to a current video. The previews represent videos from a set of videos that is currently in focus. If the user switches focus from one set of videos to another set of videos, the previews are updated to correspond to the videos in the second set of videos. The user can also browse through the previews in order to select another video for playback.

NL C 2007887

Dit octrooi is verleend ongeacht het bijgevoegde resultaat van het onderzoek naar de stand van de techniek en schriftelijke opinie. Het octrooischrift wijkt af van de oorspronkelijk ingediende stukken. Alle ingediende stukken kunnen bij NL Octrooi Centrum worden ingezien.

CONTEXTUAL VIDEO BROWSING

TECHNICAL FIELD

[0001] The present disclosure generally relates to video browsing, and specifically to browsing sets of videos that are contextually related to a current video.

BACKGROUND

[0002] For current video hosting websites, such as YOUTUBE, the user interface includes a video display window where a currently selected video is presented, along with various means by which one or more other videos may be selected for viewing. In interfaces that are designed for display on a computer screen, the user can select another video by selecting a hyperlink or a thumbnail image that is displayed in a sidebar or by entering a keyword search for a new video.

[0003] However, the arrangement, appearance, and behavior of these user interfaces are designed to be used by a user of a conventional computer having input devices such as a mouse, which allows the user to easily move a cursor to any point on the display device, and keyboard which allows for direct alphanumeric input. As such, these user interfaces become difficult to navigate when applied to televisions or other devices where the user input device provide very limited input controls. For instance, many televisions are controlled with a simple directional input device that has only up, down, left, and right buttons or keys, such as a directional pad ("D-pad") or the "arrow keys" on a keyboard. Such devices do not provide for full, arbitrary placement control over a cursor, but instead only allow for limited vertical and horizontal movements.

[0004] The problem is compounded as the number of available videos increases. Many video hosting websites host millions of videos, thereby making it all the more important to provide a user interface by which the user can easily browse and search for content of interest. Additionally, many users can now very easily view video content from video hosting sites on televisions and other display systems operated by D-pad type controllers, rather than with a computer mouse and keyboard. Users continue to expect to have the ease of use they are famil-

iar with from their computer experiences in the different system environment of a television or similar display device.

SUMMARY

5 **[0005]** The present disclosure relates to computer-implemented methods, systems and computer program products for contextual browsing of videos that are related to a currently playing video. In one embodiment, a current video is displayed to a user via a user interface. To assist the user in finding
10 another video to watch, one or more sets of videos that are related to the current video are identified, and a set of controls are displayed for navigating through the sets of related videos. The controls include labels, where each label is associated with one of the sets of related videos. In one embodi-
15 ment, the labels are arranged linearly in a first direction of the user interface (e.g. along a vertical axis) according to the relevancy between the sets of videos and the current video, such that labels for less relevant video sets are displayed at increasing distances from the label for the most
20 relevant video set.

[0006] The linear arrangement of labels allows the user to switch a focus from one label to another label to navigate between different sets of videos, using just two directional controls (e.g., up/down) on a simple directional controller.
25 As the user moves the focus to a label or between labels, previews for the set of videos represented by the focused label are displayed. These previews are preferably arranged linearly along a second direction of the user interface that is perpendicular to the first direction (e.g., along a horizontal
30 axis). This allows the user to traverse through the previews again using just two directional controls (e.g., left/right). Using the displayed previews, the user can browse through the videos in the focused set of videos and select one of the videos for playback. Through the use of these unique browsing
35 controls that present video sets based on their relevancy to the current video, the user can quickly identify another video of interest in a short amount of time.

[0007] 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.

5

BRIEF DESCRIPTION OF THE DRAWINGS

[0008] Figure (FIG.) 1 is a block diagram of the system architecture in accordance with one embodiment.

10

[0009] FIGS. 2A, 2B, 2C and 2D illustrate a user interface for contextual browsing of videos in accordance with one embodiment.

[0010] FIG. 3 illustrates a schematic view of a user interface for contextual browsing of videos in accordance with one embodiment.

15

[0011] FIG. 4 illustrates a state diagram for contextual browsing of videos in accordance with one embodiment.

20

[0012] The figures depict a preferred embodiment of the present disclosure 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 described herein.

DETAILED DESCRIPTION

25

[0013] FIG. 1 describes a system architecture of a video hosting system in accordance with one embodiment. The system includes clients 123 that communicate with a video hosting server 100 via a network 121. Generally, the video hosting server 100 is responsible for hosting uploaded media items for display to users of the clients 123. In the context of the system, the video hosting server 100 is specifically responsible for generating and providing to a video player in a client device instructions and data from which the video player can display a user interface as described herein. Alternatively, in other embodiments, the video player in the client can be responsible for generating the user interface. It should be noted that videos are just one example of media items, and in other embodiments, the principles described herein are applicable to any type of media items (e.g., pictures, music, etc).

35

[0014] As shown in FIG. 1, the video hosting server 100 comprises a front end interface 101, an ingest module 103, a video search module 105, a video serving module 107, a video set module 109, a video controls module 111, a video access log 115, a user database 117, and a video database 119. Many conventional features, such as firewalls, load balancers, application servers, failover servers, site management tools and so forth are not shown so as not to obscure the features of the system.

[0015] The video hosting server 100 can be accessed through various means. A suitable website for implementation of the video hosting server 100 is the YOUTUBE™ website, found at www.youtube.com; other video hosting sites are known as well, and can be adapted to operate according to the teaching disclosed herein. It will be understood that the term "website" represents any computer system adapted to serve content using any internetworking protocols, and is not intended to be limited to content uploaded or downloaded via the Internet or the HTTP protocol. In general, functions described in one embodiment as being performed on the server 100 side can also be performed on the client 123 side in other embodiments if appropriate. In addition, the functionality attributed to a particular component can be performed by different or multiple components operating together. The server 100 can also be accessed as a service provided to other systems or devices through appropriate application programming interfaces, and thus is not limited to use in websites only.

[0016] The video hosting server 100 is implemented as server program executing on server-class computer comprising a CPU, memory, network interface, peripheral interfaces, and other well known components. The computers themselves run an operating system such as LINUX, Microsoft Windows, or Mac OS X, have generally high performance CPUs, 2G or more of memory, and 1TB or more of disk storage. Of course, other types of computers can be used, and it is expected that as more powerful computers are developed in the future, they can be configured in accordance with the teachings here. The functionality implemented by any of the elements can be provided from computer program products (e.g., as computer executable instructions)

that are stored in tangible computer-readable storage mediums (e.g., RAM, hard disk, or optical/magnetic media).

[0017] In one embodiment a client 123 executes a video player 125 and can connect to the video hosting server 100 to view media content, herein referred to as a "video," "video content" or "video items" via a network 121. Network 121 is typically the internet, but can also be any network, including but not limited to any combination of a LAN, a MAN, a WAN, a mobile, wired or wireless network, a private network, or a virtual private network. While only a two clients 123 and video players 125 are shown, it is understood that very large numbers (e.g., millions) of clients are supported and can be in communication with the video hosting server 100 at any time.

[0018] In one embodiment, the client 123 may include a variety of different computing devices, some examples of which are televisions, set top boxes, video game consoles, personal computers, digital assistants, personal digital assistants, cellular phones, mobile phones, smart phones, tablet computers and laptop computers. The client 123 may be adapted to work with various input devices, such examples of which are a mouse, keyboard, touch-screen, or remote control.

[0019] The video player 125 includes any application that is adapted for playing and browsing videos stored on the video hosting server 100. The player can be adapted to play videos in various formats, such as Adobe Flash™, WebM, H.264, DivX, FFMPEG, and the like. In one embodiment, the video player 125 may be a standalone program that is specifically dedicated for video playback (e.g., a dedicated program in a set top box).

In other embodiments, the video player 125 is a plug-in or add on to a web browser that allows users of client 123 to access web pages on the World Wide Web. Suitable applications are Microsoft Internet Explorer, Netscape Navigator, Mozilla Firefox, Apple Safari or any application adapted to allow access to web pages on the World Wide Web.

[0020] The video player 125 allows the user of client 123 to search for and browse through videos hosted by the video hosting server 100 via a user interface provided by the front end interface 101, and loaded into the video player 125. By load-

ing the user interface in the video player 125, a user can find and play a video. The user can then select the next video for playback by browsing through different sets of videos that are related to the current video. Alternatively, the user interface can be native to the video player 125, and then populated with its content in the manner as described below.

5 **[0021]** Users of the client devices 123 can also search for videos on the user interface based on keywords, tags or other metadata. In one embodiment, these requests are received as queries in the user interface provided by the front end interface 101 and are provided to the video search module 105, which searches the video database 119 for videos that satisfy the queries. The video search module 105 supports searching on any fielded data for video content, including its title, description, tags, author, category, comment, and so forth. The video serving module 107 retrieves videos selected by the video search module 105 from the video database 119, which is formatted into a file and transmitted to the client 123 for playback in the user interface.

20 **[0022]** In some embodiments, the video player 125 also allows users to upload videos to the video hosting server 100 via the user interface provided by the front end interface 101. In one embodiment, the uploaded videos provided by users are processed by an ingest module 103 at the video hosting server 100. The ingest module 103 processes the video for storage in the video database 119. The processing can include appropriate transformations for the video content. For example, video processing can include format conversion (transcoding), compression, metadata tagging, and other data processing. The uploaded videos can include, for example, video, audio or a combination of video and audio. The ingest module 103 processes the videos in order to standardize content for playback to users of client 123. Once uploaded videos have been formatted, the ingest module 109 stores the video into the database 119. An uploaded video is associated with the content publisher (i.e., the user that uploaded the video) and the content publisher's account record is updated in the user database 117 as needed.

[0023] The video database 119 is used to store videos. The video database 119 stores videos and associated metadata provided by their respective content owner who provided the videos. Each stored video is assigned a video identifier (ID) when it is processed by the ingest module 109. The videos have metadata associated with each video such as a video ID, artist, video title, label, genre, time length, and optionally geo-restrictions that can be used for data collection or content blocking on a geographic basis.

[0024] The user database 117 stores a record of all users viewing videos provided by the video hosting server 100. In one embodiment, the user database 117 stores a record of all users who have registered an account with the video hosting server 100. Each registered account includes at least a user name and e-mail address associated with the account. A registered account may also include information about the user associated with the account such as their name, hobbies, uploaded videos, favorite videos, and/or playlists of videos created by the users.

[0025] For each user, the video access log 115 tracks the user's interactions with videos, from which relationships between videos can be determined. Each entry in the video access log 160 identifies a video being accessed, a time of access, a user ID if available, data identifying the type of interaction with the video. Interaction types can include any user interaction, such as playing, saving, rating, sharing, pausing, rewinding, and forwarding a video. The video access log may also maintain information on videos users choose to watch as measured by click thru rates between videos and/or percentage of a video watched. For example, the click thru rate between a first and second video may be the rate at which users choose (e.g., via a user input selection, such as a click) to view the second video, when presented with the first video. For example, a click thru rate can be computed as the number of times users choose to view the second video divided by the number of times the users are presented with the option of playing the second video while the first video is being played. Alternatively, the click thru rate may be simply computed as the number of times a video is selected for display

divided by the number of times the video is presented as available for displaying (e.g., a thumbnail or a link). This information as well as other types of information may be used to determine what videos are presented to a user.

5 **[0026]** The video set module 109 accesses the video access log 115, the user database 117 and/or video database 119 to identify sets of videos that are related to a video that is currently being played ("current video") in the user interface. A set of videos refers to any grouping of videos from
10 the video database 119. For example, a set of videos may include:

[0027] (1) videos sharing common metadata (e.g., artist, genre, title, etc.);

[0028] (2) videos uploaded by the same user;

15 **[0029]** (3) videos in a user-created playlist;

[0030] (4) videos having some semantic or statistical relationship with respect to the current video, such as having a high click thru or co-occurrence rate (e.g., how often two videos are watched together within some set amount of time)
20 with the current video.

[0031] The video set module 109 uses the information in the video access log 115 to determine the contextual relevance between the current video and various set of videos. For a given set of videos, contextual relevance is a measure of the likelihood (or alternatively probability) that a user watching the
25 current video would next watch a video in the given set of videos.

[0032] In one embodiment, individual relevancy scores are computed for each video in a video set as a function of user viewing patterns. For example, relevancy for a given video can be computed as a function of click thru rates for the current video and a given video, co-occurrence rates for the current video and the given video, how long users watch the given video for after viewing the current video, etc. An overall
30 relevancy score is then computed for a set of videos as a function of (e.g., an average) its individual relevancy scores. In other embodiments, other factors may be considered in determining the overall relevancy score for a set of videos, such as whether the current video is also in the set of
35

videos or whether the current video shares similar metadata with the set of videos. Sets having an overall relevancy score that is higher than some threshold are determined to be related to the current video. Sets having higher overall relevancy scores are determined to be more relevant than sets having lower overall relevancy scores.

5 [0033] The video controls module 111 manages the generation and display of contextual browsing controls for browsing of videos sets identified by the set module 109 as being related to a current videos, and provides the relevant instructions and data to the video player for displaying the controls on the client device. The video controls module 111 receives from the client 123 user inputs with respect to the browsing controls, and updates the labels on the controls, and the video set that is presented in conjunction with the controls, and provides the updated information to front end interface 101 for presentation to the client 123.

10 [0034] The browsing controls allow a user to switch between sets of videos that are related to the current video and to browse within a set of videos in order to select a new video to watch. In one embodiment, the video sets are organized according to their relevancy to the current video. The video controls module 111 receives user inputs from the client 123 via the front end interface 101 for switching between sets of videos, browsing within a set of videos, or selecting a video for playback. The user inputs are generally originated from a directional controller type of device, such as a D-pad controller, arrow keys on a keyboard, or the like, which have the semantics of "up," "down," "left", "right", and "select," though of course they can be transmitted as coded input signals. In response, the video controls module 111 updates the labels of controls in the user interface and the video set that is presented in conjunction with the controls in accordance with the user input. In some embodiments, the functionality of the video controls module 111 may be implemented in the client 123 instead of the server 100.

20 [0035] The contextual controls allow for a simple and intuitive method of selecting an additional video to watch. Because video sets, as represented by the labeled controls, are pre-

sented in arrangement based on their relevancy to the current video, the user is more likely to find a video that he/she is interested in within a short amount of time. Also, because a user can interact with the controls with a simple set of inputs (e.g., up, down, left, right, select), the interface is particularly beneficial when used in conjunction with devices that have constrained inputs. The contextual controls are explained in greater detail in conjunction with FIGS. 2B, 2C, 2D, and 3.

5
10 **[0036]** FIGS. 2A-2D illustrate a user interface 200 for contextual browsing of videos in accordance with an embodiment. In FIG. 2A, the user interface 200 includes a video 205 that has been selected by a user for playback, here for example a video of a car race. The video 205 may have been selected, for example, from a list of videos displayed to the user after entering a search query or from the browsing controls shown in FIG. 2B-2D. The video 205 occupies the entire interface 200 as it is being played. Alternatively, the video 205 may occupy only a portion of the interface 200 instead of the entire interface 200.

15
20 **[0037]** While watching the video 205 the user provides an input to activate the display of the controls, for example the "up" key. In FIG. 2B, contextual browsing controls 210 are overlaid onto the current video 205 in the user interface 200 in response to this input. Alternatively, the browsing controls may also be displayed when the video 205 finishes playing.

25
30 **[0038]** The browsing controls 210 include pivot controls 215 and set controls 220. The pivot controls 215 allow the user to switch focus between sets of videos, and the set controls 220 allow the user to switch focus between videos within the set of videos that is currently in focus. For example, the user may use the up or down buttons of a remote control to switch between sets of videos and the right or left buttons to switch between videos in a set. Once the user locates a video that he or she wants to watch, the user selects the video, for example with an enter button. The currently playing video 205 is then replaced with the selected video and the controls 210 are removed from the display.

35

[0039] The pivot controls 215 are comprised of individual labels (e.g., 215-1, 215-2, 215-3). Each label is a visual element that represents a set of videos. The labels 215 are displayed along a first axis (e.g., a vertical axis) in a linear arrangement. In some embodiments, a label includes text that summarizes or explains the nature of the video set represented by the label. For example, label 215-1 represents the set of videos from a user created playlist called "My Favorites." Label 215-2 represents the set of videos uploaded by user "Jim2000." Label 215-3 represents the set of videos in the topic of "Auto Racing." The label that has the focus is displayed in a visually distinctive manner (e.g., with highlighting, color, size, motion, outlining, etc). For example, label 215-1 has the focus and is shown with a larger and darker boarder. The user can move the focus between labels 215 using, for example, up and down inputs.

[0040] The set controls 220 are comprised of a plurality of preview cells (e.g., 220-1, 220-2, 220-3, 220-4) for the set of videos that is currently in focus, which can be referred to for convenience as the current playlist. For example, the label 215-1 for the playlist "My Favorites" is in focus, and so the set controls 220 include previews for videos in the playlist "My Favorites" to allow the user to browse among the videos in this playlist. The preview cells 220 are displayed in a linear arrangement along a second axis of the user interface 200 (e.g. a horizontal axis) that is perpendicular to the first axis.

[0041] Each preview cell is a visual element that represents a video. The preview cells can include, for example, still images or clips from the videos represented by the cell, and/or textual information about the videos (e.g., title, length, rating, total views, user comments, other metadata) to facilitate the user's decision of which video to watch. In one embodiment, the preview cell that is in focus can include textual information about its video, whereas the preview cells that are not in focus do not include the textual information. As shown, one of the previews cells 220-2 is displayed in a visually distinctive manner to indicate the cell 220 having the focus, and the user can move the focus between preview

cells 220 using left and right inputs. In one embodiment, each time the set controls 220 are initially displayed, the preview cell 220 for the current video 205 is given the focus. Thus, in the FIG. 2B, preview cell 220-2, which represents the current video 250, is initially given the focus.

[0042] As previously discussed, the contextual relevance between the current video 205 and each set of videos is determined by the video set module 109, and represented by a relevancy score. The labels 215 for the sets of videos are organized in the interface 200 according to the relevance of the associated video sets to the current video 205, for example, in order from less relevant sets to more relevant sets. For example, the set of videos in the "My Favorite" playlist (label 215-1) is more relevant than the set of videos that are uploaded by user "Jim2000" (label 215-2), which in turn is more relevant than the videos in the topic of "Auto Racing" (label 215-3). Organizing the video sets according to their relevancy is beneficial for presenting the user with videos that he/she is likely to be interested in watching. In one embodiment, when the controls are first displayed, the set of videos with the highest relevancy (e.g., label 215-1) is located closest to the set controls 220, and less relevant sets of videos are located at increasing distances from the set controls 220. Focus is also placed on the set of videos that has the highest relevancy (e.g., label 215-1).

[0043] As shown, the contextual browsing controls 210 are overlaid onto the current video 205 while the current video 205 continues to play in the background. This allows the user to continue watching the current video 205 while browsing for a new video to watch. Alternatively, the current video 205 may be placed in a smaller window and displayed alongside the browsing controls 210, the current video 205 may be hidden from view, and/or the current video 205 may be paused while the browsing controls 210 are displayed.

[0044] FIG. 2C illustrates how the set controls 220 can be used to browse through a set of videos according to an embodiment. In FIG. 2C, a user input has been received for shifting the focus from preview cell 220-2 to preview cell 220-3, for example a "right" input. In response, the focus is moved from

the preview cell 220-2 to preview cell 220-3, which is displayed in a visually distinctive manner to show that focus has shifted. The current video 205 and the labels in the pivot controls 215 remain relatively unchanged as focus is shifted from one preview cell to the next, thereby maintaining the context of the video sets relative to the current video 205. Note that in FIG. 2B and 2C, in the preview cell with the focus there is displayed additional contextual information, such as the title of the video and the number of times it has been viewed.

[0045] FIG. 2D illustrates how the pivot controls 215 can be used to switch between sets of videos. In FIG. 2D, a user input has been received for shifting the focus from label 215-1 to label 215-2, for example, and "up" input. In response, the focus is moved from the label 215-1 to label 215-2, which is displayed in a visually distinctive manner to show that focus has shifted

[0046] In one embodiment, the pivot controls 215 scroll vertically across the user interface 200 as focus changes from one label to another. For example, now that label 215-2 is in focus, all the labels are shifted towards the bottom of the user interface 200. If label 215-3 were to be selected for focus, the labels would again scroll toward the bottom of the screen. A new label 215-4, which previously did not fit in the user interface 200, is now displayed in the user interface that represents a set of videos in the genre of "Sports." This shifting of labels can be repeated any number of times, and with each shift a new label 215 is displayed that corresponds to another set of videos. Thus, the user can continue to traverse "up" through any number of sets of videos. In some embodiments, the pivot controls 200 do not scroll, and are instead relatively stationary in the user interface 200 as the user switches from one set to another.

[0047] The set controls 220 are updated to correspond to the set of videos that is currently in focus. Thus, in Fig. 2D, the set controls 220 are updated to show previews for videos in the set of videos uploaded by user "Jim2000." Because the pivot controls 215 were scrolled downwards, label 215-1, which was previously in focus, is now located directly below and ad-

5 adjacent to the pivot controls 220. Label 215-2, which is now in focus, is located directly above and adjacent to the set controls 220. In other words, each time the user selects another label 215, the browsing controls 210 are updated so that the previews 220 are displayed between the label 215 with the focus and the previously focused label 215. As shown in Fig. 2D, the position of the set controls 220 is relatively stationary as the user switches focus from one label 215 to another. In some embodiments, the position of the set controls 220 may
10 move up or down in the user interface to track the position of the label 215 with focus. Keeping the set controls 220 adjacent to the label in focus helps the user maintain a sense of the relationship between the label that is in focus and the previews shown in the set controls 220.

15 **[0048]** FIG. 3 illustrates a schematic view of a user interface for contextual browsing of videos according to an embodiment. As shown, a current video 205 is playing in the background of the user interface 200. The pivot controls 215 include five labels that are distributed in a vertical direction
20 of the user interface along the vertical axis 305. In other embodiments, a fewer or greater number of labels can be displayed. The labels are organized along the vertical axis 305 according to the relevancy between the current video 205 and the video sets represented by the labels. For example, more
25 relevant video sets may be represented by labels placed at the bottom of the user interface 200, and less relevant video sets may be represented by labels placed at the top of the user interface 200.

30 **[0049]** Because the focus of the pivot controls 215 is on the label for video set n 215-5, the set controls 220 also correspond to videos from video set n. Previews of the videos in video set n are displayed in the set controls 220. The previews 220 are distributed in a horizontal direction of the user interface along a horizontal axis 310. The pivot controls
35 220 are located directly beneath the label 215-5 with focus. Generally speaking, the location of the pivot controls 220 is linked to the location of the label with focus, which in this example is label 215-5. The horizontal axis 310 and vertical

axis 305 are perpendicular to each other. Note that the axes may or may not be located at the center of the user interface.

[0050] FIG. 4 is a state diagram for contextual browsing of videos according to an embodiment. Initially, a video is being played 510 to a user via a user interface. A user input for displaying browsing controls is received 511, wherein the user input is a directional input or a select/enter input. In response, the set module 109 identifies video sets that are related to the currently playing video and determines 515 the relevancy of these sets to the currently playing video. In some embodiment, the relevancy of the video sets may be determined in step 510 before any user input is received.

[0051] Once the relevancy of video sets is determined 515, contextual browsing controls are displayed 520 to the user via the user interface. The user interface may be similar to that shown in FIGS. 2B, 2C, 2D or FIG. 3. The contextual browsing controls allow a user to switch between sets of videos and to browse videos within each set of videos. The video sets are organized in the user interface according to their relevancy to expedite the process of finding a new video to watch.

[0052] The system then waits 525 for a user input. If a user input is received 531 for changing focus from one video set to another or for changing focus within a video set, the browsing controls are updated 530 in the user interface 525 in the manner described with respect to FIGS. 2A-2D, and the system waits 525 for another user input. If a user input is received 526 that selects a video for playback, the system hides the browsing controls and plays 510 the video in the user interface, and the process starts anew.

[0053] The apparatus described above for performing the described operations 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 in the computer. Such a computer program may be stored in a 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 cir-

5 cuits (ASICs), or any type of media suitable for storing elec-
tronic instructions, and each coupled to a computer system
bus. Furthermore, the computers referred to in the specifica-
tion may include a single processor or may be architectures
employing multiple processor designs for increased computing
capability.

10 **[0054]** Some portions of above description describe the em-
bodiments in terms of algorithms and symbolic representations
of operations on information. These algorithmic descriptions
and representations are commonly used by those skilled in the
data processing arts to convey the substance of their work ef-
fectively to others skilled in the art. These operations,
while described functionally, computationally, or logically,
are understood to be implemented by computer programs or
15 equivalent electrical circuits, microcode, or the like. Fur-
thermore, it has also proven convenient at times, to refer to
these arrangements of operations as modules, without loss of
generality. The described operations and their associated mod-
ules may be embodied in software, firmware, hardware, or any
20 combinations thereof.

25 **[0055]** As used herein any reference to "one embodiment" or
"an embodiment" means that a particular element, feature,
structure, or characteristic described in connection with the
embodiment is included in at least one embodiment. The appear-
ances of the phrase "in one embodiment" in various places in
the specification are not necessarily all referring to the
same embodiment.

30 **[0056]** As used herein, the terms "comprises," "comprising,"
"includes," "including," "has," "having" or any other varia-
tion thereof, are intended to cover a non-exclusive inclusion.
For example, a process, method, article, or apparatus that
comprises a list of elements is not necessarily limited to
only those elements but may include other elements not ex-
pressly listed or inherent to such process, method, article,
35 or apparatus. Further, unless expressly stated to the con-
trary, "or" refers to an inclusive or and not to an exclusive
or. For example, a condition A or B is satisfied by any one of
the following: A is true (or present) and B is false (or not

present), A is false (or not present) and B is true (or present), and both A and B are true (or present).

[0057] In addition, use of the "a" or "an" are employed to describe elements and components of the embodiments herein.

5 This is done merely for convenience and to give a general sense of the disclosed embodiments. This description should be read to include one or at least one and the singular also includes the plural unless it is obvious that it is meant otherwise. Similarly, the terms "first" and "second" are generally
10 used herein to allow clear referencing of different elements of a same class (e.g., a "first label" as simply distinguished from "a second label"), rather than designating an ordering, ranking, or placement of the first item relative to the second item.

15 **[0058]** Upon reading this disclosure, those of skill in the art will appreciate still additional alternative structural and functional designs for a system and a process for browsing video sets through the disclosed principles herein. Thus, while particular embodiments and applications have been illustrated and described, it is to be understood that the disclosed embodiments are not limited to the precise construction
20 and components disclosed herein. Various modifications, changes and variations, which will be apparent to those skilled in the art, may be made in the arrangement, operation and details of the method and apparatus disclosed herein without departing from the spirit and scope defined in the appended claims.
25

CONCLUSIES

1. Systeem voor het onderzoeken van video's via een gebruikersinterface, welk systeem omvat:

een computerprocessor; en

5 een computer-leesbaar opslagmedium welke instructies opslaat voor het besturen van de processor voor het uitvoeren van een werkwijze omvattende:

het weergeven in de gebruikersinterface van een meertal labels gedistribueerd langs een eerste richting van de gebruikersinterface;

10 waarin het meertal labels correspondeert met een meertal video-verzamelingen, waarbij iedere videoverzameling is gerelateerd aan een actuele video, en waarin de labels geordend zijn langs de eerste richting volgens een relevantie van iedere videoverzameling met betrekking tot de actuele video, waarin gefocusseerd is op een eerste label van het meertal labels, 15 en het eerste label correspondeert met een eerste videoverzameling; en

het verschaffen voor weergave in de gebruikersinterface van een meertal video-previews gedistribueerd langs een tweede 20 richting van de gebruikersinterface die loodrecht is op de eerste richting,

waarin de previews corresponderen met de video's uit de eerste videoverzameling, en

25 waarin de previews tegelijkertijd worden weergegeven met de labels.

2. Systeem volgens conclusie 1, verder omvattende het in reactie op een gebruikersinvoer voor het verplaatsen van de focus onder de labels van het eerste label naar een tweede label 30 corresponderende met een tweede videoverzameling, updaten van de videopreviews zodat deze corresponderen met video's van de tweede videoverzameling in plaats van de video's van de eerste videoverzameling.

3. Systeem volgens conclusie 2, waarin het eerste label gelokaliseerd is aan een eerste zijde van de videopre- 35 views, wanneer de focus ligt op het eerste label en waarin het eerste label gelokaliseerd is op een tweede zijde van de vi-

deopreviews tegenover de eerste zijde, wanneer de focus ligt op het tweede label.

5 4. Systeem volgens conclusie 2, waarin de videopreviews gelokaliseerd zijn tussen het eerste label en de tweede label wanneer de focus rust op de tweede label.

5. Systeem volgens conclusie 1, waarin focus onder het meertal previews ligt op een eerste preview corresponderend met een eerste video van de eerste videoverzameling.

10 6. Systeem volgens conclusie 5, waarin de eerste preview tekstuele informatie omvat die geassocieerd is met de eerste video van de eerste videoverzameling.

15 7. Systeem volgens conclusie 5, verder omvattende het in reactie op een gebruikersinvoer die de eerste preview selecteert wanneer de focus ligt op de eerste preview, weergeven van de eerste video van de eerste videoverzameling in de gebruikersinterface.

8. Systeem volgens conclusie 1, waarin een actuele video een actueel spelende video omvat.

20 9. Systeem volgens conclusie 1, waarin een label een tekstuele beschrijving van de videoverzameling die correspondeert met het label omvat.

10. Systeem volgens conclusie 1, waarin een preview een stilstaand beeld omvat van de video die correspondeert met de preview.

25 11. Systeem volgens conclusie 1, waarin een relevantie van een videoverzameling wordt bepaald gebaseerd op kijkpatronen tussen de actuele video en video's in de videoverzameling.

200

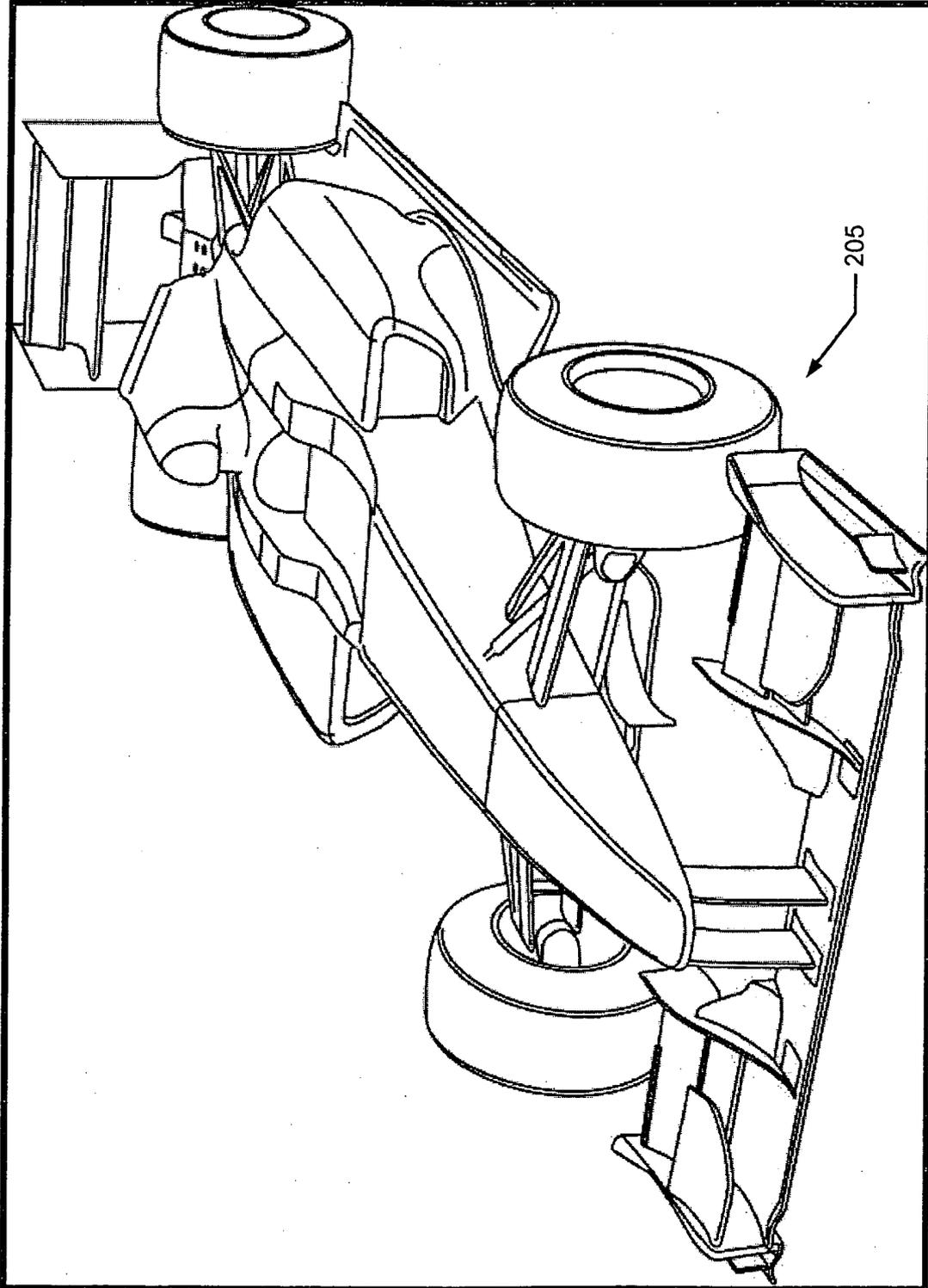


FIG. 2A

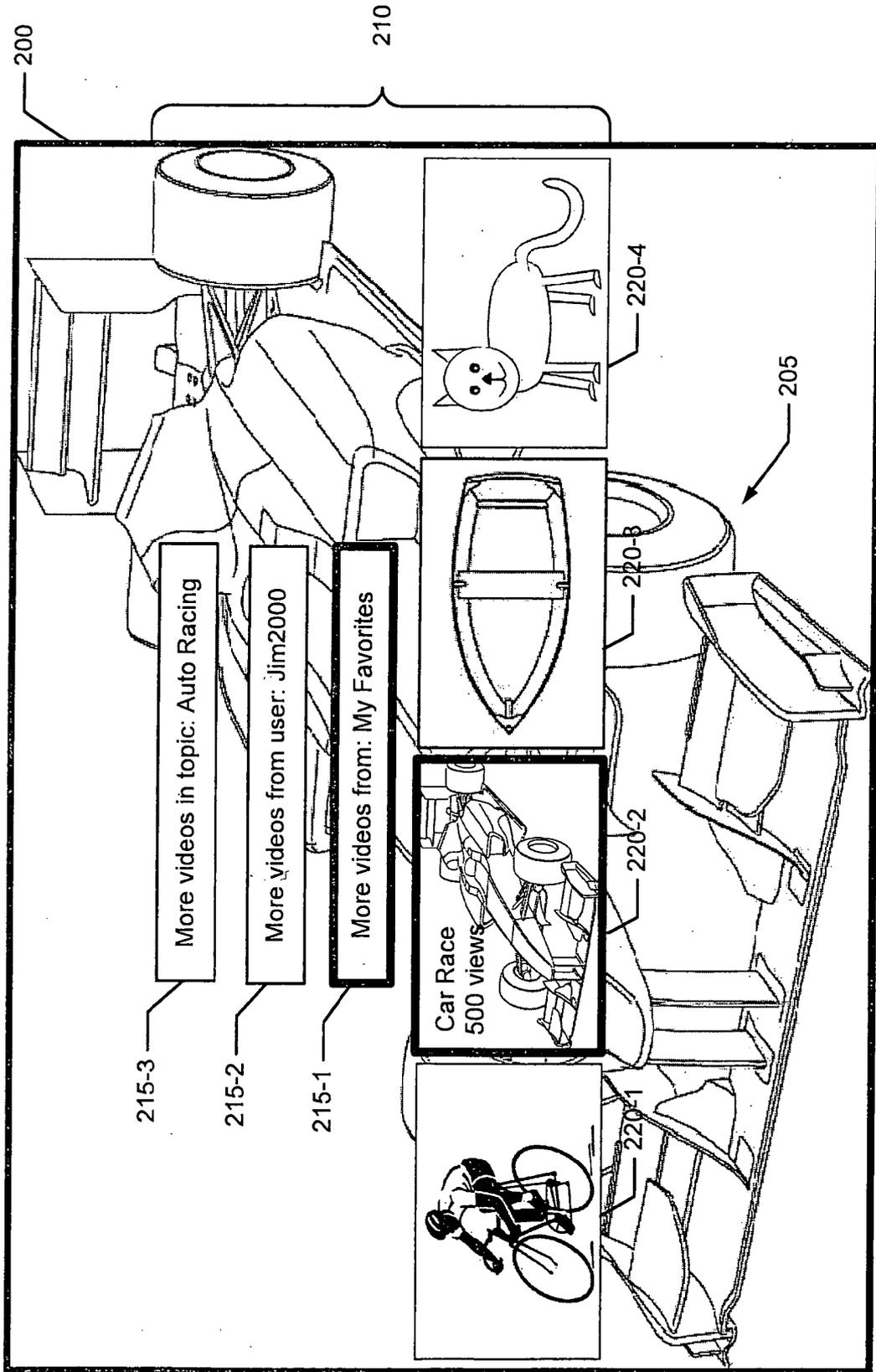


FIG. 2B

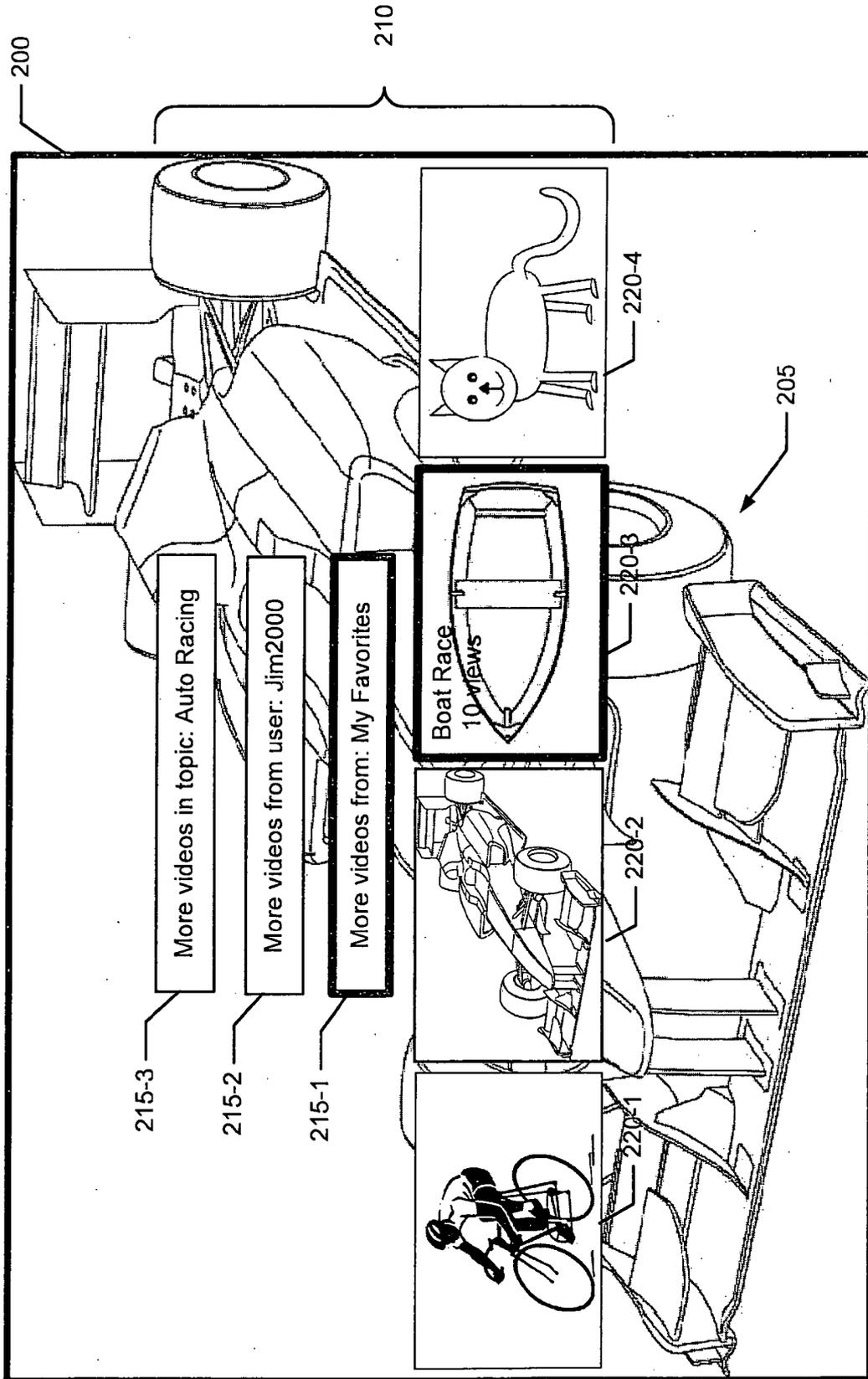


FIG. 2C

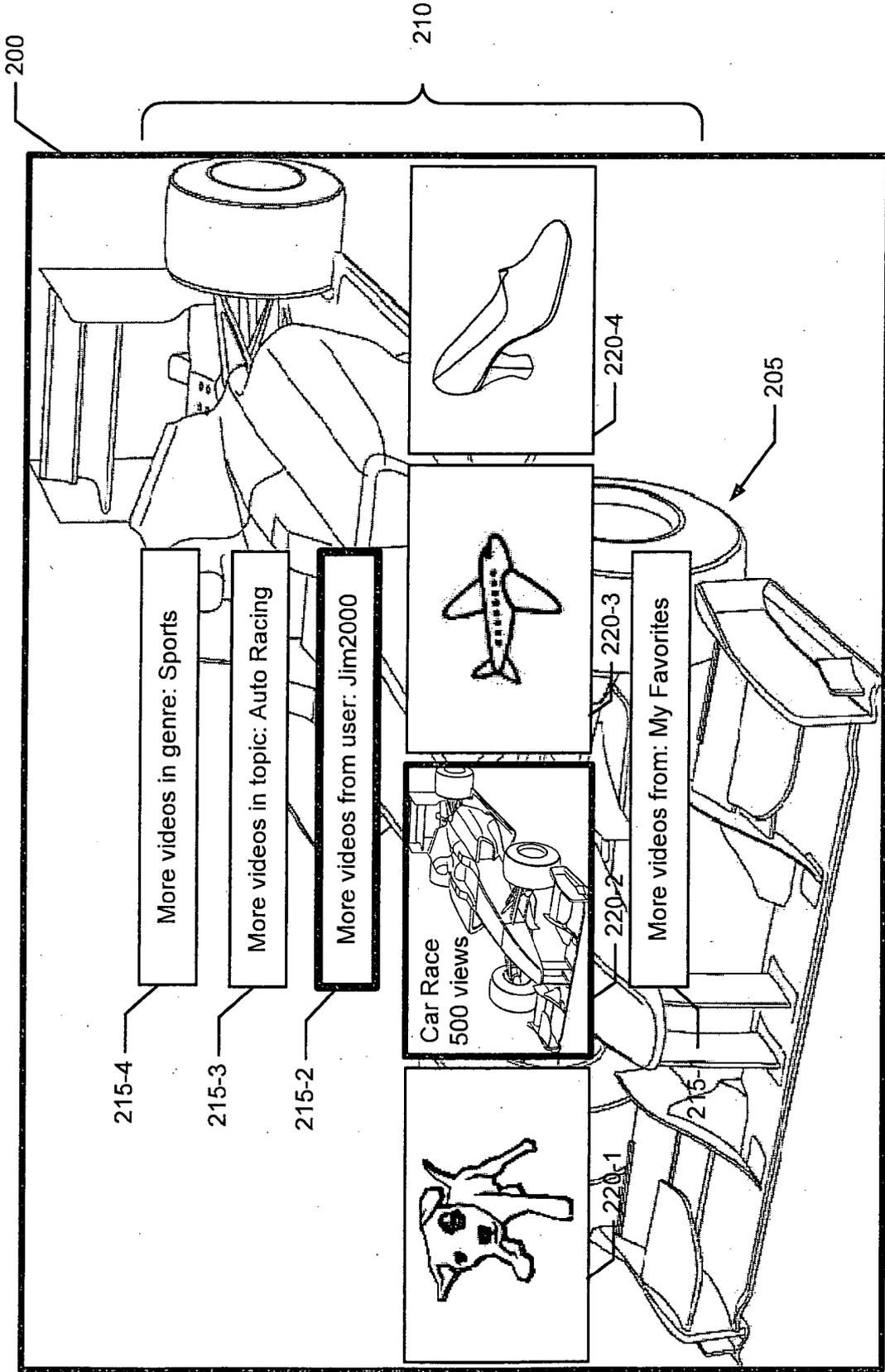


FIG. 2D

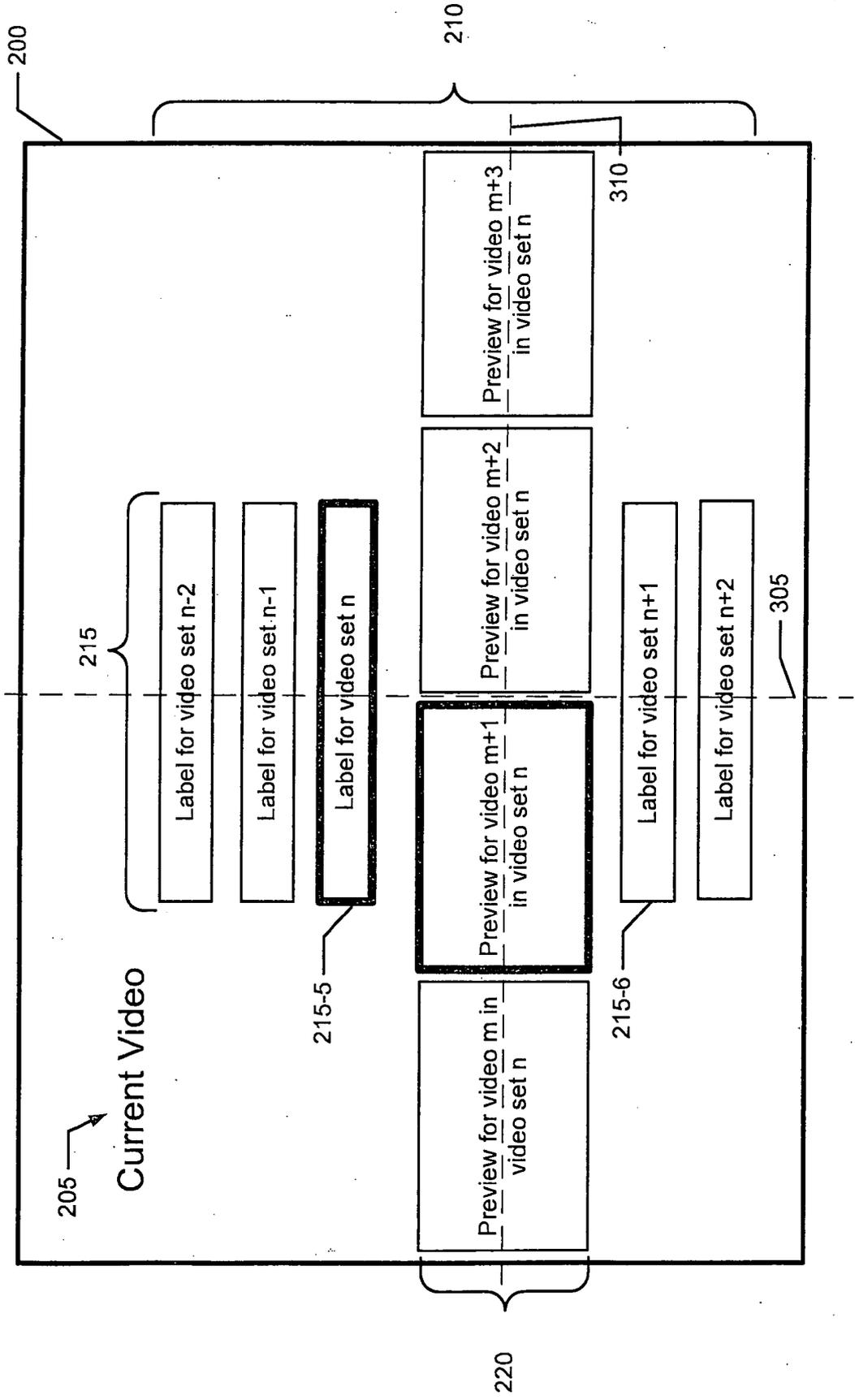


FIG. 3

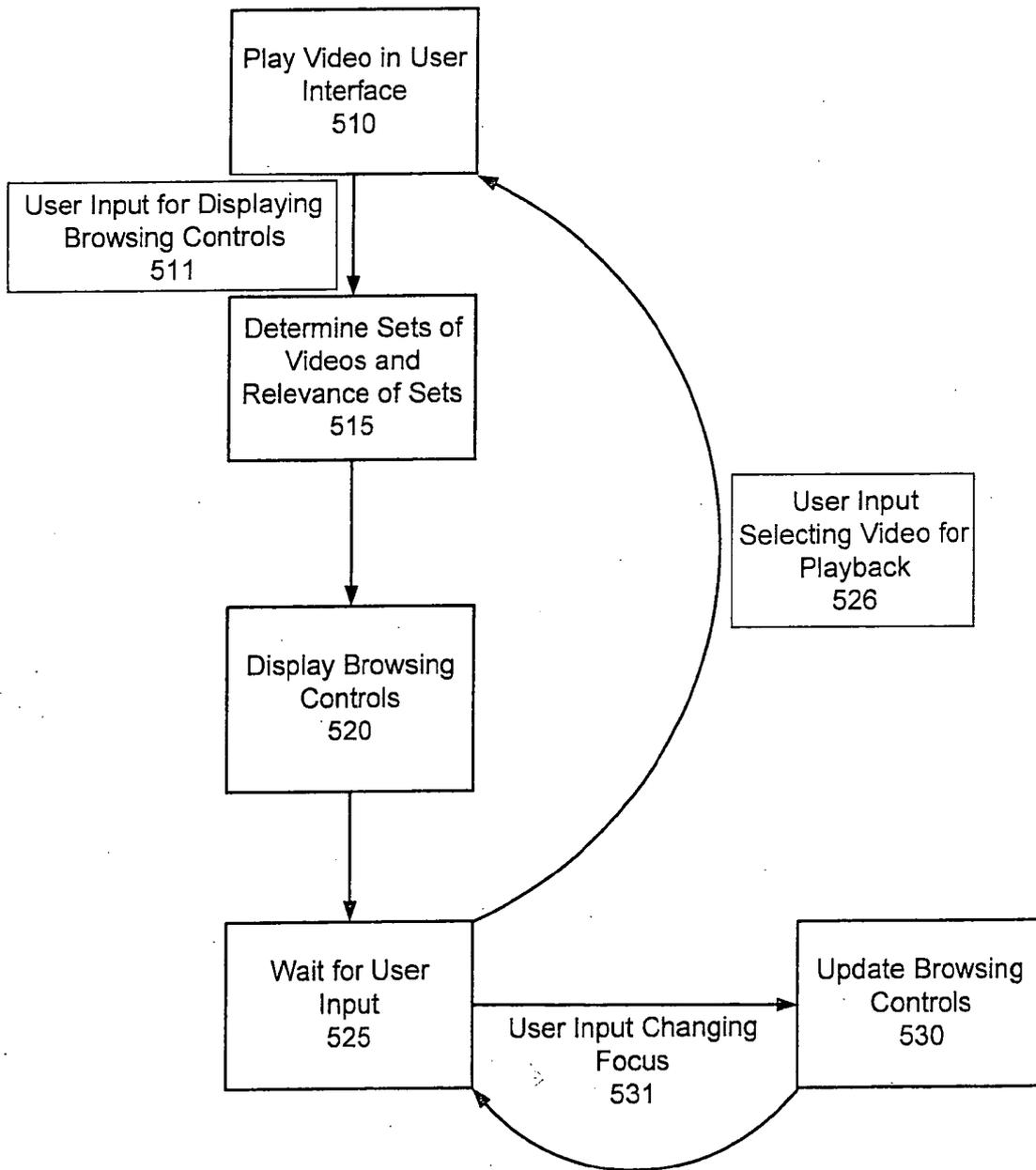


FIG. 4



ONDERZOEKSRAPPORT

BETREFFENDE HET RESULTAAT VAN HET ONDERZOEK NAAR DE STAND VAN DE TECHNIEK

RELEVANTE LITERATUUR			
Categorie ¹	Literatuur met, voor zover nodig, aanduiding van tekstgedeelten of figuren.	Van belang voor conclusie(s) nr.	Classificatie (IPC)
X	US 2007/174790 A1 (JING FENG [CN] ET AL) 26 juli 2007 (2007-07-26) * alinea [0023] - [0044]; figuren 1-15 *	1-11	INV. G06F17/30 G06F3/048
A	JAMES DAVIDSON ET AL: "The YouTube video recommendation system", PROCEEDINGS OF THE FOURTH ACM CONFERENCE ON RECOMMENDER SYSTEMS, RECSYS '10, 30 september 2010 (2010-09-30), bladzijde 293, XP55021645, New York, New York, USA DOI: 10.1145/1864708.1864770 ISBN: 978-1-60-558906-0 * bladzijden 294-295, alinea 2; figuur 1 *	1,8,10, 11	
A	US 2008/074399 A1 (LEE CHANG SUB [KR]) 27 maart 2008 (2008-03-27) * alinea [0037]; figuur 6c *	3,4	
A	ALEX HINDLE ET AL: "Clustering Web video search results based on integration of multiple features", WORLD WIDE WEB ; INTERNET AND WEB INFORMATION SYSTEMS, KLUWER ACADEMIC PUBLISHERS, DO, deel 14, nr. 1, 3 augustus 2010 (2010-08-03), bladzijden 53-73, XP019870643, ISSN: 1573-1413, DOI: 10.1007/S11280-010-0097-X * bladzijden 59-68, alinea 4; figuren 2-5 *	1-11	
			Onderzochte gebieden van de techniek
			G06F H04N
Indien gewijzigde conclusies zijn ingediend, heeft dit rapport betrekking op de conclusies ingediend op:			
Plaats van onderzoek:	Datum waarop het onderzoek werd voltooid:	Bevoegd ambtenaar:	
München	20 maart 2012	Legrand, J	

¹ CATEGORIE VAN DE VERMELDE LITERATUUR

X: de conclusie wordt als niet nieuw of niet inventief beschouwd ten opzichte van deze literatuur
Y: de conclusie wordt als niet inventief beschouwd ten opzichte van de combinatie van deze literatuur met andere geciteerde literatuur van dezelfde categorie, waarbij de combinatie voor de vakman voor de hand liggend wordt geacht
A: niet tot de categorie X of Y behorende literatuur die de stand van de techniek beschrijft
O: niet-schriftelijke stand van de techniek
P: tussen de voorrangsdatum en de indieningsdatum gepubliceerde literatuur

T: na de indieningsdatum of de voorrangsdatum gepubliceerde literatuur die niet bezwarend is voor de octrooiaanvraag, maar wordt vermeld ter verheldering van de theorie of het principe dat ten grondslag ligt aan de uitvinding
E: eerdere octrooi(aanvraag), gepubliceerd op of na de indieningsdatum, waarin dezelfde uitvinding wordt beschreven
D: in de octrooiaanvraag vermeld
L: om andere redenen vermelde literatuur
&: lid van dezelfde octrooifamilie of overeenkomstige octrooipublicatie

**AANHANGSEL BEHORENDE BIJ HET RAPPORT BETREFFENDE
HET ONDERZOEK NAAR DE STAND VAN DE TECHNIEK,
UITGEVOERD IN DE OCTROOIAANVRAGE NR.**

NO 138057
NL 2007887

Het aanhangsel bevat een opgave van elders gepubliceerde octrooiaanvragen of octrooien (zogenaamde leden van dezelfde octroofamilie), die overeenkomen met octrooischriften genoemd in het rapport.

De opgave is samengesteld aan de hand van gegevens uit het computerbestand van het Europees Octrooibureau per

De juistheid en volledigheid van deze opgave wordt noch door het Europees Octrooibureau, noch door het Bureau voor de Industriële eigendom gegarandeerd; de gegevens worden verstrekt voor informatiedoeleinden.

20-03-2012

In het rapport genoemd octrooigeschrift	Datum van publicatie	Overeenkomend(e) geschrift(en)	Datum van publicatie
US 2007174790 A1	26-07-2007	US 2007174790 A1	26-07-2007
		US 2010088647 A1	08-04-2010
US 2008074399 A1	27-03-2008	TW 200825886 A	16-06-2008
		US 2008074399 A1	27-03-2008
		US 2011025632 A1	03-02-2011



Agentschap NL
Ministerie van Economische Zaken,
Landbouw en Innovatie

SCHRIFTELIJKE OPINIE

DOSSIER NUMMER NO138057	INDIENINGSDATUM 29.11.2011	VOORRANGSDATUM 09.05.2011	AANVRAAGNUMMER NL2007887
CLASSIFICATIE INV. G06F17/30 G06F3/048			
AANVRAGER Google Inc.			

Deze schriftelijke opinie bevat een toelichting op de volgende onderdelen:

- Onderdeel I Basis van de schriftelijke opinie
- Onderdeel II Voorrang
- Onderdeel III Vaststelling nieuwheid, inventiviteit en industriële toepasbaarheid niet mogelijk
- Onderdeel IV De aanvraag heeft betrekking op meer dan één uitvinding
- Onderdeel V Gemotiveerde verklaring ten aanzien van nieuwheid, inventiviteit en industriële toepasbaarheid
- Onderdeel VI Andere geciteerde documenten
- Onderdeel VII Overige gebreken
- Onderdeel VIII Overige opmerkingen

	DE BEVOEGDE AMBTENAAR Legrand, J
--	-------------------------------------

SCHRIFTELIJKE OPINIE

Aanvraag nr.:
NL2007887

Onderdeel I Basis van de Schriftelijke Opinie

1. Deze schriftelijke opinie is opgesteld op basis van de meest recente conclusies ingediend voor aanvang van het onderzoek.
2. Met betrekking tot **nucleotide en/of aminozuur sequenties** die genoemd worden in de aanvraag en relevant zijn voor de uitvinding zoals beschreven in de conclusies, is dit onderzoek gedaan op basis van:
 - a. type materiaal:
 - sequentie opsomming
 - tabel met betrekking tot de sequentie lijst
 - b. vorm van het materiaal:
 - op papier
 - in elektronische vorm
 - c. moment van indiening/aanlevering:
 - opgenomen in de aanvraag zoals ingediend
 - samen met de aanvraag elektronisch ingediend
 - later aangeleverd voor het onderzoek
3. In geval er meer dan één versie of kopie van een sequentie opsomming of tabel met betrekking op een sequentie is ingediend of aangeleverd, zijn de benodigde verklaringen ingediend dat de informatie in de latere of additionele kopieën identiek is aan de aanvraag zoals ingediend of niet meer informatie bevatten dan de aanvraag zoals oorspronkelijk werd ingediend.
4. Overige opmerkingen:

SCHRIFTELIJKE OPINIE

Aanvraag nr.:
NL2007887

Onderdeel V Gemotiveerde verklaring ten aanzien van nieuwheid, inventiviteit en industriële toepasbaarheid

1. Verklaring

Nieuwheid	Ja: Conclusies 1-11 Nee: Conclusies
Inventiviteit	Ja: Conclusies Nee: Conclusies 1-11
Industriële toepasbaarheid	Ja: Conclusies 1-11 Nee: Conclusies

2. Citaties en toelichting:

Zie aparte bladzijde

Onderdeel VIII Overige opmerkingen

De volgende opmerkingen met betrekking tot de duidelijkheid van de conclusies, beschrijving, en figuren, of met betrekking tot de vraag of de conclusies nawerkbaar zijn, worden gemaakt:

Zie aparte bladzijde

Regarding Item V

Prior Art

- 1 Reference is made to the following documents:
- D1 US 2007/174790 A1 (JING FENG [CN] ET AL) 26 juli 2007 (2007-07-26)
- D2 JAMES DAVIDSON ET AL: "The YouTube video recommendation system",
PROCEEDINGS OF THE FOURTH ACM CONFERENCE ON RECOMMENDER SYSTEMS, RECSYS '10, 30 september 2010 (2010-09-30), bladzijde 293, XP55021645,
New York, New York, USA
DOI: 10.1145/1864708.1864770
ISBN: 978-1-60-558906-0
- D3 US 2008/074399 A1 (LEE CHANG SUB [KR]) 27 maart 2008 (2008-03-27)
- D4 ALEX HINDLE ET AL: "Clustering Web video search results based on integration of multiple features",
WORLD WIDE WEB ; INTERNET AND WEB INFORMATION SYSTEMS, KLUWER ACADEMIC PUBLISHERS, DO,
deel 14, nr. 1, 3 augustus 2010 (2010-08-03), bladzijden 53-73, XP019870643,
ISSN: 1573-1413, DOI: 10.1007/S11280-010-0097-X

Lack of inventive step

- 2 The present application does not meet the criteria of patentability, because the subject-matter of claim 1 does not involve an inventive step.
- 2.1 Document D1 is regarded as being the prior art closest to the subject-matter of claim 1, and discloses:
- Systeem voor het onderzoeken van video's via een gebruikersinterface (§45: "The user interface system may be used to present content of various types, such as photographs, drawings, artwork, **videos**, music, and so on"), welk systeem omvat:
- een computerprocessor (§35: "central processing unit"); en

een computer-leesbaar opslagmedium welke instructies opslaat voor het besturen van de processor (§ 35: "the memory and storage devices are computer-readable media that may contain instructions that implement the user interface system") voor het uitvoeren van een werkwijze omvattende:

het weergeven in de gebruikersinterface van een meertal labels (§23: "each cluster of images includes a cluster name", §27: "the name of the cluster as a link", fig. 2(225)) gedistribueerd langs een eerste richting van de gebruikersinterface (see fig. 2);

waarin het meertal labels correspondeert met een meertal videoverzamelingen (§23, fig. 2), waarbij iedere videoverzameling is gerelateerd aan een zoekopdracht, en waarin de labels geordend zijn langs de eerste richting volgens een relevantie van iedere videoverzameling met betrekking tot de actuele zoekopdracht (§23: "the cluster areas of the cluster panel may be ordered based on a relevance score of the images of the cluster to the image query"), waarin gefocuseerd is op een eerste label van het meertal labels (§26: "when the lucky view is selected, the user interface system may display the thumbnails of the images of the **first cluster** of the search result"), en het eerste label correspondeert met een eerste videoverzameling (§27, fig. 2(231-238)); en

het verschaffen voor weergave in de gebruikersinterface van een meertal videopreviews (§4: "thumbnails") gedistribueerd langs een tweede richting van de gebruikersinterface die loodrecht is op de eerste richting (see fig. 2: the thumbnails of the individual videos are displayed horizontally), waarin de previews corresponderen met de video's uit de eerste videoverzameling (§23: "the cluster area for a cluster may include [...] mini-thumbnails of some of the images of the cluster."), en waarin de previews tegelijkertijd worden weergegeven met de labels (see fig. 2).

- 2.2 The subject-matter of claim 1 therefore differs from the disclosure of D1 in that, in claim 1 each video set is related to a current video and the labels are ordered according to a relevancy to the current video, whereas in D1 each video set is related to a current text query and the labels are ordered according to a relevancy to the current query (§ 26, fig. 1).
- 2.3 The problem to be solved by the subject-matter of claim 1 may therefore be regarded as how to provide an alternative way of querying the system of D1.
- 2.4 Using a "current video" as input to a video search, i.e. performing a "query by example" belongs to the common general knowledge in the field of video retrieval (see for example D2, p. 294, §2.2: "Related videos"). Therefore, the

skilled person would apply the teaching of D1 to a video recommendation system, where the retrieved video sets relates to a current (seed) video, without the need for any inventive activity.

- 3 Dependent claims 2-11 do not contain any additional features which, in combination with the features of any claim to which it/they refer, meet the requirements of inventive step, the reasons being as follows:
- 3.1 Claim 2: see D1, §30-31, fig. 5-6.
- 3.2 Claims 3, 4: it is generally considered that the arrangement of the menu items on the screen, is not a technical aspect of a menu driven control system. Moreover, the claimed arrangement corresponds to one of the design choices available to the skilled person (see for example D3, §37, fig. 6(c)).
- 3.3 Claim 5: in the system of D1, the user can implicitly put the focus on the first video from the first video set (§25: "the user interface system displays a detail view of an image in the view panel when a user selects an image (e.g., by selecting a thumbnail from the view panel or mini-thumbnail from a thumbnail list)").
- 3.4 Claim 6: see fig. 2, a title and image size is displayed next to each preview.
- 3.5 Claim 7: when the system of D1 displays videos (§45), it would be obvious to play a video when selected.
- 3.6 Claim 8: see D2, p. 294, right column: "In this context, we define similar videos as those that a user is likely to watch after having watched the given seed video v".
- 3.7 Claim 9: see D1, § 23: "cluster name", fig. 2(225).
- 3.8 Claim 10: both D1 (fig. 2) and D2 (§2.5, fig. 1) disclose still images as preview.
- 3.9 Claim 11: see D2, §2.2.

Regarding Item VIII

Lack of clarity

- 4 In addition, it should be noted that claims 3-5 are not clear:
- 4.1 Claims 3, 4: the expressions "on a first side" and "on a second side" are unclear, because it is not defined in which space and according to which orientation they apply.

- 4.2 Claim 5: claim 1 claims that the "focus among the plurality of labels is on a first label", whereas claim 5 claims "focus among the plurality of previews in on the first preview". It is unclear how the focus can be both on the first label and on the first preview.