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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
17 June 2021 (17.06.2021)

WIPO | PCT

(10) International Publication Number
WO 2021/119150 A1

(51) International Patent Classification:
G10L 15/22 (2006.01)

(21) International Application Number:
PCT/US2020/064048

(22) International Filing Date:
09 December 2020 (09.12.2020)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
16/709,734 10 December 2019 (10.12.2019) US

(71) Applicant: **ROVI GUIDES, INC.** [US/US]; 2160 Gold Street, San Jose, California 95002 (US).

(72) Inventors: **ROBERT JOSE, Jeffry Copps**; C30, Lakshmi Narasimha Nilaya, 1st Main, 4th Cross, Bangalore 560017 (IN). **AHER, Ankur**; 11, Giriraj Krupa C.H.S., Rambaug Lane No. 2, Maharashtra, Kalyan 421301 (IN).

(74) Agent: **GUILIANO, Joseph M.** et al.; HALEY GUILIANO LLP, 75 Broad Street, Suite 1000, New York, New York 10004 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report (Art. 21(3))*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

(54) Title: SYSTEMS AND METHODS FOR PROVIDING VOICE COMMAND RECOMMENDATIONS

(57) Abstract: The system provides a voice command recommendation to a user to avoid a non-voice command. The system determines a command that is expected to be received, and generates a voice command recommendation that corresponds to the predicted command. The predicted command can be based on the user's behavior, a plurality of users' behavior, environmental circumstances such as a phone call ring, or a combination thereof. The system may access one or more databases to determine the predicted command. The voice command recommendation may include a displayed notification that describes the recommended voice command, and exemplary voice inputs that are recognized. The system also activates an audio interface, such as a microphone, that is configured to receive a voice input. If the system receives a recognizable voice input at the audio interface that corresponds to the recommendation, the system performs the predicted command in response to receiving the voice input.

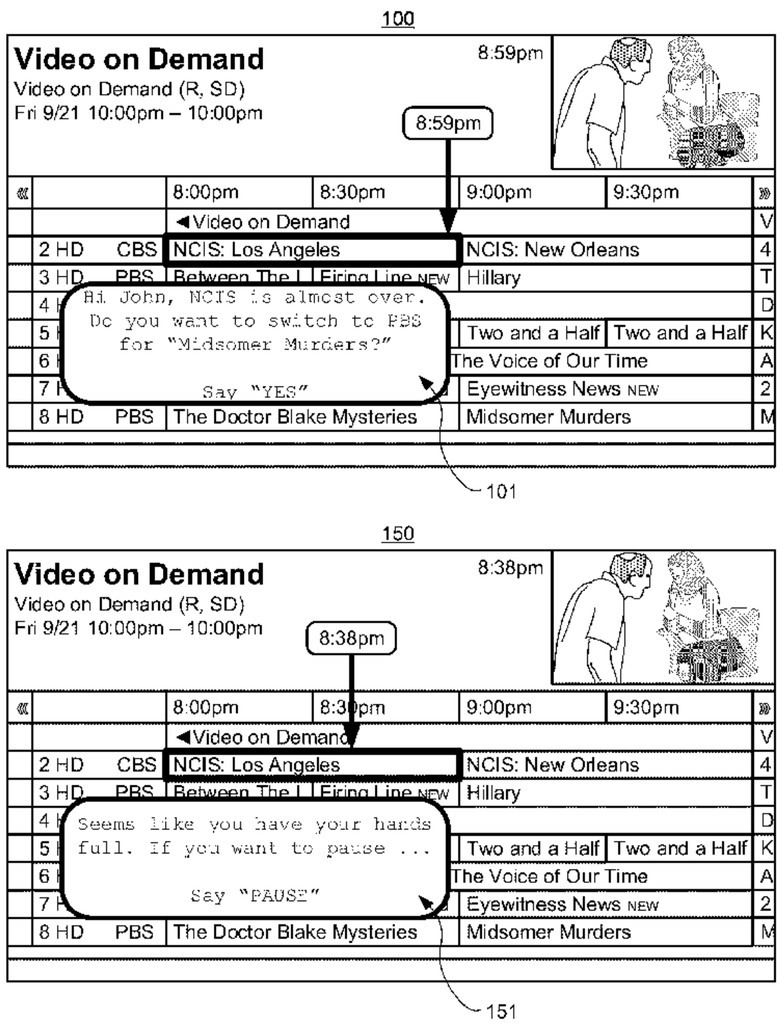


FIG. 1

SYSTEMS AND METHODS FOR PROVIDING VOICE COMMAND RECOMMENDATIONS

Background

[0001] The present disclosure relates to systems for generating voice command recommendations, and more particularly to systems for generating voice command recommendations based on predicted interactions.

Summary

[0002] Conversation systems have made interactions between entertainment systems and users easy and comfortable. However, in some circumstances and for various reasons, users still choose alternate ways to interact with the system other than voice commands. For example, users typically interact with the system using remote controllers and it may be difficult to change such habits of users. In a further example, users might not be aware that a particular interactive function can be achieved with the system by using voice commands. In a further example, users may not know or remember the exact command to execute a particular interactive function.

[0003] The present disclosure describes systems and methods for identifying situations which are most convenient for a user to use voice commands instead of using remote controllers, and then suggest appropriate voice commands to the user. In some embodiments, the system predicts that the user will be imminently interacting with the system, based on history of the user's behavior and profile information, based on environmental cues, or based on any other suitable inputs or combination of inputs. The system constructs a suggestion for a voice command, displays the suggestion to the user, and prepares an audio interface or device, such as a microphone, to receive audio input. In

response, the user can provide a voice command. If the voice command corresponds to the suggestion, the system will perform the interactive function.

Brief Description of the Drawings

[0004] The above and other objects and advantages of the disclosure will be apparent upon consideration of the following detailed description, taken in conjunction with the accompanying drawings, in which like reference characters refer to like parts throughout, and in which:

[0005] FIG. 1 shows two illustrative displays including generated voice commands, in accordance with some embodiments of the present disclosure;

[0006] FIG. 2 is a block diagram of an illustrative system for generating voice command recommendations, in accordance with some embodiments of the present disclosure;

[0007] FIG. 3 is a block diagram of illustrative user devices, in accordance with some embodiments of the present disclosure;

[0008] FIG. 4 is a flowchart of an illustrative process for generating a voice command recommendation, in accordance with some embodiments of the present disclosure; and

[0009] FIG. 5 shows a flowchart of an illustrative process for managing voice command interactions, in accordance with some embodiments of the present disclosure.

Detailed Description

[0010] The present disclosure describes systems and methods for identifying the most convenient time and circumstance for a user to use voice commands instead of using remote controllers, and then suggest appropriate voice commands to the user. This functionality will help the system to be more assistive to the user by requiring minimal efforts for any interactive function. This functionality will help the user to understand the usage and advantages of a conversation system.

[0011] In illustrative example, a user is currently watching some latest sitcom episode on channel HBO. As part of a usual routine, after the sitcom ends, the user changes the channel to Sky News to watch a news program. Generally, user changes the channel using a remote control. If the remote control is not in the vicinity of the user and is out of reach for the user to grab it, the system of the present disclosure identifies this situation. The system predicts based on the user's history that the user wants to change the channel to Sky News as the sitcom ends. The system also detects that user cannot easily grab the TV remote control to

change the channel. In response, the system identifies this opportunity to notify user of the voice command that can be used to change the channel. The system displays a notification on the television display such as "To change channel just say – Tune to Sky News." The system also activates a microphone to receive the voice command. Accordingly, if the user speaks a voice command as suggested by the notification, the system proceeds to change the channel.

[0012] FIG. 1 shows illustrative displays 100 and 150, which include generated voice commands, in accordance with some embodiments of the present disclosure. For example, displays 100 and 150 may be generated on a television screen, computer monitor, user device screen (e.g., a touchscreen of a mobile phone), or any other suitable display device. A system configured to generate displays 100 and 150 identifies optimal times to provide a voice command suggestion as a notification to the user. The system may take into account that the desired interactive function will be difficult to achieve for the user by any other means apart from voice command, or otherwise that a voice command suggestion may be useful to the user. Thus, the system improves the chances that user will use a voice command.

[0013] Display 100, as illustrated includes a grid of program identifiers with the vertical axis defined by channel (e.g., or station) and the horizontal axis defined by predetermined display time. For example, time slots ranging from 8:00pm to 10:00pm are illustrated for seven channels. The current clock time as illustrated in display 100 is 8:59pm, with the cursor highlighting the program "NCIS: Los Angeles." The system identifies that it is almost 9:00pm on Friday, and that the user typically switches to channel 8 to watch "Midsomer Murders" on PBS. The system generates notification 101, which indicates to the user the predicted command that they are likely to provide, and also includes an indication "Say, YES" of how the user can provide that command as a voice command. When the system generates notification 101, the system also activates a microphone to receive the voice command should the user choose to provide voice input. If the user responds with a voice command of "Yes," the system will perform the channel change command (e.g., changing the channel to PBS). If the user does not respond (e.g., within some predetermined time) with the recommended voice command, the system does not perform any action (e.g., and may cease displaying notification 101 after some predetermined time period).

[0014] Display 150, as illustrated includes a grid of program identifiers with the vertical axis defined by channel (e.g., or station) and the horizontal axis defined by

predetermined display time. For example, time slots ranging from 8:00pm to 10:00pm are illustrated for seven channels. The current clock time as illustrated in display 150 is 8:38pm, with the cursor highlighting the program "NCIS: Los Angeles." At 8:38pm, the system identifies an environmental event such as a phone call, smart doorbell ring, or reduced proximity of the user to the system or remote control. The system generates notification 151, which a predicted command of pausing the program, and also includes an indication "Say, YES" of how the user can provide that command as a voice command. When the system generates notification 151, the system also activates a microphone to receive the voice command should the user choose to provide voice input. If the user responds with a voice command of "Yes," the system will pause playback of the current program. If the user does not respond (e.g., within some predetermined time) with the recommended voice command, the system does not perform any action (e.g., and may cease displaying notification 151 after some predetermined time period, or repeat the notification at a later time).

[0015] FIG. 2 is a block diagram of illustrative system 200 for generating voice command recommendations, in accordance with some embodiments of the present disclosure. System 200 includes interactive system 220, condition information 270, clock/time information 280, user information 290, and display device 250. Interactive system 220, as illustrated, includes display generator 221, non-voice input interface 222, voice input interface 223, command predictor 224, recommendation generator 225, and condition analyzer 226. Interactive system 220 detects the interactive function needed, or likely needed, by the user and notifies the user of a corresponding voice command. The notification may be displayed on the television screen (e.g., a primary screen) or on a peripheral device such as a mobile phone screen (e.g., a secondary screen). In an illustrative example, system 200 is configured to (1) predict an imminent user-interaction with non-voice input interface 222 or voice input interface 223 using command predictor 224, (2) identify that the predicted interaction will be difficult to achieve using an interactive device (say remote controller) for the user at that time using condition analyzer 226, and (3) notify the user with a displayed recommendation of an appropriate voice command for the interactive function by displaying it on the system screen using display generator 221. In a further illustrative example, system 200 is configured to (1) predict an imminent user-interaction with non-voice input interface 222 based on user behavior using command predictor 224, (2) generate a recommendation for a voice command to achieve the user-interaction using

recommendation generator 225, (3) notify the user with a displayed recommendation of an appropriate voice command for the interactive function by displaying it on the system screen using display generator 221, and (4) activate voice input interface 223 to receive voice input. In a further illustrative example, system 200 is configured to (1) predict an imminent action to be performed based on one or more users' behavior using command predictor 224, (2) generate a recommendation using recommendation generator 225 for a voice command to achieve the action using display generator 221, and (4) activate voice input interface 223 to receive voice input.

[0016] Condition information 270 includes information about conditions that may affect a user's use of voice commands. Condition information 270 may include current information, historical information, statistical information, or any combination thereof for one or more users. For example, condition information may include position information (e.g., historical position information, current position information, or both), events (e.g., logs or a current instance of a phone ringing, a doorbell ringing, a device interaction, an alarm), any other suitable information, or any combination thereof. For example, condition information 270 may include information about when content is consumed on a mobile phone or other interactive device. In some embodiments, condition information 270 includes information about other users in proximity to the first user. In some embodiments, condition information 270 is included as part of interactive system 220. For example, information of condition information 270 may be stored in memory storage of interactive system 220, and accordingly may be retrieved based on one or more criteria. In some embodiments, as illustrated, condition information 270 need not be included as part of interactive system 220 and be implemented on hardware of a separate device or system that is communicatively coupled (e.g., via a wired, wireless, or optical network) to interactive system 220.

[0017] Clock/time information 280 includes information such as, for example, a current clock time, a clock time associated with a user command or other input, a display time of a program (e.g., start time, end time, or duration), program schedules, any other temporal information, or any combination thereof. Clock/time information 280 is stored in any suitable memory on any suitable device or combination of devices. In some embodiments, for example, clock/time information 280 is determined by control circuitry. The time may be based on an oscillator circuit (e.g., a "real time clock" or RTC), a time communicated from a network entity (e.g., to calibrate a software clock), a clock implemented in hardware and/or software, any other suitable time reference, or

any combination thereof. In some embodiments, clock/time information 280 includes historical temporal information for one or more users such as, for example, time a user command was inputted (e.g., times when a currently viewed channel is changed), viewing times (e.g., start time, end time, duration), any other suitable temporal information, or any combination thereof. For example, clock/time information 280, or a portion thereof, may be stored as a time stamp or other suitable time indicator accompanying data stored in user information 290. Accordingly, user information 290 may be combined with at least some of clock/time information 280.

[0018] User information 290 may include user identification information (e.g., name, an identifier, address, contact information), user command history such as previous voice commands or previous remote control commands, user preferences (e.g., search settings, favorite media content, temporal viewing preferences), user likes/dislikes (e.g., user-inputted information, user selections), user voice data (e.g., audio samples, signatures, speech patterns, or files for identifying the user's voice), any other suitable information about a user, or any combination thereof. User information 290 include any suitable information for predicting a voice command that may be useful to the user based on their past behavior and profile information. In some embodiments, user information 290 may be included in one or more databases. In some embodiments, user information 290 include statistical information for a plurality of users (e.g., search histories, content consumption histories, consumption patterns), a plurality of entities (e.g., content associated with entities, metadata, static types), or both. For example, user information 290 may include information about a plurality of users' behavior, selected commands, voice commands, remote control commands, any other suitable information, or any combination thereof. In some embodiments, user information 290 is included as part of interactive system 220. For example, information of user information 290 may be stored in memory storage of interactive system 220, and accordingly may be retrieved based on one or more criteria. In some embodiments, as illustrated, user information 290 need not be included as part of interactive system 220 and be implemented on hardware of a separate device or system that is communicatively coupled (e.g., via a wired, wireless, or optical network) to interactive system 220. In an illustrative example, user information 290 may store any suitable interaction information, as illustrated in Table 1:

Table 1 - Record of various interactive functions used by user

INTERACTIVE FUNCTION	NO. OF TIMES USING REMOTE	LAST USAGE TIMESTAMP VIA HANDHELD REMOTE	NO. OF TIMES USING VOICE	LAST USAGE TIMESTAMP VIA VOICE COMMAND
Channel change	1945	2019-08-06	0	---
Mute	122	2019-07-04	3	2018-11-02 01:34:07
Open App	411	2019-06-01	7	2019-01-06 14:19:55
Volume change	1545	2019-08-06	1	2018-08-02 22:08:37
...	...	...	...	...

As illustrated in Table 1, interactions such as channel changes, volume adjustments, and other commands may be stored, along with frequency and date/time, in user information 290. In some embodiments, user information 290 stores records for other interactive devices apart from a TV remote such as, for example, a game controller/console, a keyboard, a trackball, a mobile phone, or any other device(s), and use the stored values to determine whether to provide a voice command notification. In some embodiments, user information 290 is combined with condition information 270 (e.g., stored in a common database(s)). For example, although not shown in Table 1, position information of a user, a device, or both, may also be included in such a record.

[0019] Display generator 221 is configured to generate a display on display device 250 (e.g., a television screen, computer monitor, or smartphone touchscreen). In some embodiments, system 200 includes or is coupled to display device 250, which is configured to provide a visual display (e.g., display 251) of information to the user. In an illustrative example, display 251 may be, but need not be, similar to displays 100 and 150 of FIG. 1. In some embodiments, display generator 221 determines display parameters to determine how to organize, arrange, configure, or otherwise present information. Display parameters may include, for example, screen size, zoom extent, pixel count or resolution, available colors or color palette, placement and properties of overlays (e.g., notification windows), user preferences, any other suitable parameters for generating a display, or any combination thereof. In some embodiments, display device 250 provides required or suggested display parameters to display generator 221. In some embodiments, interactive system 220 (e.g., or display generator 221 thereof) includes a software driver or other software having display parameters for generating a display (e.g., on display device 250 or any other suitable display device). For example,

interactive system 220 or display generator 221 thereof may include a graphics card or other suitable hardware controller that determines display parameters based on properties of the display device (e.g., display device 250) and the information to be displayed on the display device (e.g., display 251 to be displayed on display device 250). In some embodiments, depending on the display type, size and layout, display generator 221 modifies a pitch, size, level of detail, or other aspect of a displayed grid. In an illustrative example, display generator 221 generates a displayed grid on display device 250. In some embodiments, interactive system 220 or display generator 221 thereof transmits video data to display device 250 to generate display 251. Display 251 may include content identifiers, content descriptions, time indicators, a cursor, a current time (e.g., a clock time), notifications, any other suitable features, or any combination thereof. For example, interactive system 220 or display generator 221 thereof may transmit a high-definition multimedia interface (HDMI) signal to display device 250, over a suitable cable or connection, to generate a grid including the episode identifiers (e.g., similar to displays 100 and 150 of FIG. 1, or display 251 of FIG. 2). In some embodiments, interactive system 220 or display generator 221 thereof may update an existing display on display device 250. For example, a first display may be presented on display device 250, and interactive system 220 or display generator 221 thereof may update the first display to a second display by transmitting a video signal having some portions updated and some portions similar or maintained as before. In some embodiments, interactive system 220 or display generator 221 thereof generates a grid for display that is configured to fit into a target space on display device 250.

[0020] Non-voice input interface 222 may include any suitable interface such as, for example, a receiver for receiving signals from a handheld remote control, hard buttons (e.g., switches, push-buttons, or any other suitable user-actuated buttons), a touchscreen (e.g., with soft buttons or other selectable or adjustable elements), a keypad (e.g., with a preset layout for entering text commands), any other suitable interface that does not process voice input, or any combination thereof. Non-voice input interface 222 requires a non-voice interaction which may include a haptic action by the user (e.g., touching an area on a screen or pressing a button).

[0021] Voice input interface 223 may include, for example, a microphone for receiving audio signals, signal processing hardware, signal processing software, speech detection software, voice identification software, any other suitable hardware or software for processing voice input, or any combination thereof. In some embodiments,

voice input interface 223 includes an audio interface that includes a microphone or other sensor that receives audio input and generates an electronic signal. In some embodiments, the audio input is received at an analog sensor, which provides an analog signal that is conditioned, sampled, and digitized to generate an audio file. The audio file may then be analyzed by interactive system 220. Voice input interface 223 may be configured to receive audio input at any time, when speech or other suitable audio signals are detected, when activated by interactive system 220 (e.g., command predictor 224, recommendation generator 225, or condition analyzer 226 thereof), in response to any other criterion, or any combination thereof. Interactive system 220 or voice input interface 223 thereof may include any suitable conditioning software or hardware for converting audio input to a stored audio file or electronic signal. For example, interactive system 220 or voice input interface 223 thereof may apply one or more filters (e.g., low-pass, high-pass, notch filters, or band-pass filters), amplifiers, decimators, or other conditionings to generate the audio file. In a further example, interactive system 220 or voice input interface 223 thereof may apply any suitable processing to a conditioned signal to generate an audio file such as compression, transformation (e.g., spectral transformation, wavelet transformation), normalization, equalization, truncation (e.g., in a time or spectral domain), any other suitable processing, or any combination thereof. Interactive system 220 or voice input interface 223 thereof may be configured to extract one or more keywords from the received voice input. For example, in some embodiments, the query application may identify words in speech, and select some of those words as keywords. Interactive system 220 or voice input interface 223 thereof may process words using any suitable word detection technique, speech detection technique, pattern recognition technique, signal processing technique, or any combination thereof. For example, interactive system 220 or voice input interface 223 thereof may compare a series of signal templates to a portion of an audio signal to find whether a match exists (e.g., whether a particular word is included in the audio signal). In a further example, interactive system 220 or voice input interface 223 thereof may store snippets (i.e., clips of short duration) of recorded audio during detected speech and process the snippets. In some embodiments, the query application stores relatively large segments of speech (e.g., more than 10 seconds) as an audio file, and processes the file. In some embodiments, interactive system 220 or voice input interface 223 thereof may process speech to detect words by using a continuous computation. For example, a wavelet transform may be performed on speech in real time, providing a continuous

computation of speech patterns (e.g., which could be compared to a reference to identify words). In some embodiments, interactive system 220 or voice input interface 223 thereof may detect words, as well as which user uttered the words (e.g., voice recognition) in accordance with the present disclosure.

[0022] Command predictor 224 is configured to predict a command that is (1) expected to be inputted based on a user's past behavior (e.g., from use information 290), (2) likely to be inputted based on a plurality of users' past behavior, (3) useful to the user based on one or more events, or (4) useful to the user as a way to gain experience with interactive system 220 using voice commands. In some embodiments, command predictor 224 retrieves the user's interaction history with interactive system 220 from user information 290 and analyzes it to identify patterns. For example, at particular times of day or night the user may watch particular channels or particular shows. The channel and time (e.g., start time, end time, duration, or combination thereof) may be stored in user information 290. In an illustrative example, for particular types of channels (e.g., music channels), the user may increase the volume as compared to the volume for watching television programs. User information 290 maintains watch history and remote control usage history for every interaction of the user with interactive system 220. In some embodiments, command predictor 224 determines a command based on input from condition analyzer 226, which is configured to identify one or more environmental or circumstance-based conditions relevant to the user. For example, during certain types of disturbances, (e.g., telephone call ring, doorbell ring, crying baby) the user might generally pause the currently displayed program, or mute the volume. The history of such interactions is stored in user information 290, and condition analyzer 226 identifies a condition that corresponds to the stored interactions. For example, condition analyzer 226 access external cues (e.g., telephone call ring, doorbell ring, or crying baby) in the room and predicts a command of use to the user. Command predictor 224 may predict any suitable type of command such as, for example, changing a channel, changing a volume, muting a volume, pausing/playing a program, rewinding/forwarding a program, exiting or stopping a program, initiating or starting a program, recording a program, displaying information about a program (e.g., a plot summary, a list of actors, ratings information), displaying information based on a program (e.g., related programs), setting an alarm, clearing an alarm, navigation-based commands (e.g., go to a billing page, go to an on-demand page), any other suitable command, or any combination thereof (e.g., both change a channel and decrease a volume). In an illustrative example, command predictor

224 may generate an interactive function prediction map for the user as illustrated in Table 2:

Table 2 – Prediction map of illustrative functions.

CONDITION	INTERACTIVE FUNCTION
After watching GOT on HBO	Change channel to Sky News
After tuning to music channel	Increase TV volume
On receiving a phone call	Pause the program, Mute the volume
On consuming content on a mobile device	Pause the program
At 9:00 pm	Open Netflix
On detecting lost proximity of user and remote	Pause the program
Lack of voice command usage	Any suitable function
Following a non-voice command	The just-inputted command
After a predetermined duration	Any suitable function
…	…

As illustrated in Table 2, a condition may include program endings or beginnings, user behavior (e.g., receiving a phone call, nitrating with a mobile device, moving around), a time (e.g., a clock time), a duration of a lack of voice command usage, use of a non-voice command, any other suitable condition, or any combination thereof. In some embodiments, interactive system 220 generates, retrieves, or otherwise uses tabulated values to determine if a voice command notification should be displayed to the user or not. For example, interactive system 220 need not generate a voice command notification for commands that are regularly used and known by the user. To illustrate, interactive system 220 may consider the number of times a voice command is used for an interactive function as well as the last usage timestamp (e.g., which may be tabulated similar to arrangement of Table 1). If both of these values lie below certain thresholds, then interactive system 220 generates a voice command notification. In some embodiments, interactive system 220 compares the usage of a remote control and usage of voice commands for one or more interactive functions. In some such embodiments, if the difference between the times of usage is greater than a threshold (e.g., with remote control usage on higher side), the system may generate a voice command notification.

[0023] Recommendation generator 225 is configured to determine and generate a recommendation for a voice command. The recommendation may be in the form of a notification, as illustrated by notifications 101 and 151 of FIG. 1. The recommendation may be based on which command is predicted by command predictor 224, which

command is the most convenient or useful for the user, or a combination thereof. For example, recommendation generator 225 may identify voice commands will be the most convenient way for the user to achieve an interactive function at that moment (e.g., as compared to a using a remote control). In some embodiments, recommendation generator 225 highlights the reason in the notification to the user for using voice command. The highlight may include adding information about the device that is out of reach for the user, the detected event, or other suitable information. For example, recommendation generator may generate a notification that includes "Cannot find TV remote? Just say 'Volume Up' to increase the volume," or "Away from keyboard? Just say 'TV mute' to mute the TV."

[0024] Condition analyzer 226 is configured to detect, determine, identify, or otherwise analyze conditions that may affect a user's use of voice commands. In some embodiments, interactive system 220 or condition analyzer 226 thereof monitors or otherwise determines the position of one or more interactive devices (e.g., a remote control, a game controller, a keyboard, a trackball, a mobile phone, or other device), the position of the user (e.g., in the room or elsewhere), or both. For example, if the distance between user and the interactive device is more than a predetermined threshold, interactive system 220 or condition analyzer 226 identifies that it will be difficult for the user to grab the interactive device (e.g., the condition is "lack of proximity"). In some embodiments, interactive system 220 uses cameras to map and determine (e.g., calculate) the position of the user and the position of the interactive devices in the room (e.g., the room wherein interactive system 220 or components thereof reside). In some embodiments, interactive system 220 or condition analyzer 226 thereof determines the user's position based on tracking one or more peripheral devices such as, for example, a smart watch, fitness tracker, a pair of headphones, any other suitable peripheral device, or any combination thereof. In some embodiments, interactive system 220 or condition analyzer 226 thereof maps interactive devices through detected signals arising from electronic circuitry of the interactive device.

[0025] Condition analyzer 226 may include any suitable sensors, signal processing circuitry, or both to detect one or more conditions. For example, condition analyzer 226 may include a system that interacts with a remote control or other interactive device to determine the interactive device's position. To illustrate, condition analyzer 226 may determine distance or position based on a signal (e.g., a latency in a returned signal, a signal strength, an attenuation of a signal, or changes thereof). In a further example, condition analyzer 226 may use image analysis to determine movement, distance, proximity, or

changes thereof. In some embodiments, an interactive device or peripheral device may detect its own position and communicate position information to interactive system 220 using any suitable communications link (e.g., wireless, near-field, or optical transmission). In some embodiments, condition analyzer 226 retrieves information from condition information 270. In some embodiments, condition analyzer 226 detects one or more conditions and generates information to transmit to condition information 270 for storage.

[0026] In some embodiments, condition information 270 includes information about other users in proximity to the first user. For example, if two users are in proximity to interactive system 220, condition analyzer 226 may identify conditions wherein one user leaves and may generate a corresponding voice command recommendation (e.g., pause the program until they return to proximity, or increase the volume to reach the greater distance). In a further example, if one user is in proximity to interactive system 220, condition analyzer 226 may identify conditions wherein another user arrives and may generate a corresponding voice command recommendation (e.g., pause the program until they get situated).

[0027] In an illustrative example, interactive system 220 may generate a voice command suggestion (e.g., using recommendation generator 225) to the user having a "yes" or "no" answer. User information 290 may include template information, the user's history of interacting with templates, or both. In some embodiments, the yes/no question will take into consideration previous template usage of the user to generate the question. For example, interactive system 220 may have three templates for changing a channel: "Go to channel <channel-name>," "Tune to channel <channel-name>," and "Jump to channel <channel-name>." Accordingly, interactive system 220 may generate different questions for same interactive function. For example, interactive system 220 may generate notifications such as "Would you like to go to channel Sky News?," "Would you like to tune to channel Sky News?," and "Would you like to jump to channel Sky News?" Based on which template the user is more familiar to considering the past query usage of the user, interactive system 220 will select the best template amongst the three options. In some embodiments, interactive system 220 may be communicatively coupled to a voice-activated home assistant to present the question to the user.

[0028] In some embodiments, interactive system 220 or a remote control includes a microphone that may not have support to be triggered by a ‘wake word’. For example, the user may press a button on the remote control to input a voice command. In some embodiments, interactive system 220 triggers the microphone (e.g., for five seconds or any

other suitable time) when it presents a suggestion to the user and is ready to receive the user's voice response. In some embodiments, microphone is not configured to be always active to receive voice input. For example, the microphone may need to be activated to receive and record a voice command. In some such circumstances, when interactive system 220 generates the voice command notification to the user, it also activates the microphone for a certain period (e.g., thirty seconds or any other suitable time) because the user is expected to use the voice command soon after receiving the notification.

[0029] FIG. 3 is a block diagram of illustrative user devices, in accordance with some embodiments of the present disclosure. User equipment system 301 may include set-top box 316 that includes, or is communicatively coupled to, display 312, audio equipment 314, and user input interface 310. In some embodiments, display 312 may include a television display or a computer display. In some embodiments, user input interface 310 is a remote-control device. Set-top box 316 may include one or more circuit boards. In some embodiments, the one or more circuit boards include processing circuitry, control circuitry, and storage (e.g., RAM, ROM, hard disk, removable disk, etc.). In some embodiments, circuit boards include an input/output path. Each one of user device 300 and user equipment system 301 may receive content and data via input/output (hereinafter "I/O") path 302. I/O path 302 may provide content and data to control circuitry 304, which includes processing circuitry 306 and storage 308. Control circuitry 304 may be used to send and receive commands, requests, and other suitable data using I/O path 302. I/O path 302 may connect control circuitry 304 (and specifically processing circuitry 306) to one or more communications paths (described below). I/O functions may be provided by one or more of these communications paths but are shown as a single path in FIG. 3 to avoid overcomplicating the drawing. While set-top box 316 is shown in FIG. 3 for illustration, any suitable computing device having processing circuitry, control circuitry, and storage may be used in accordance with the present disclosure. For example, set-top box 316 may be replaced by, or complemented by, a personal computer (e.g., a notebook, a laptop, a desktop), a network-based server hosting a user-accessible client device, a non-user-owned device, any other suitable device, or any combination thereof.

[0030] Control circuitry 304 may be based on any suitable processing circuitry such as processing circuitry 306. As referred to herein, processing circuitry should be understood to mean circuitry based on one or more microprocessors, microcontrollers, digital signal processors, programmable logic devices, field-programmable gate arrays

(FPGAs), application-specific integrated circuits (ASICs), etc., and may include a multi-core processor (e.g., dual-core, quad-core, hexa-core, or any suitable number of cores) or supercomputer. In some embodiments, processing circuitry is distributed across multiple separate processors or processing units, for example, multiple of the same type of processing units (e.g., two Intel Core i7 processors) or multiple different processors (e.g., an Intel Core i5 processor and an Intel Core i7 processor). In some embodiments, control circuitry 304 executes instructions for an application stored in memory (e.g., storage 308). Specifically, control circuitry 304 may be instructed by the application to perform the functions discussed above and below. For example, the application may provide instructions to control circuitry 304 to generate the media guidance displays. In some implementations, any action performed by control circuitry 304 may be based on instructions received from the application.

[0031] In some client/server-based embodiments, control circuitry 304 includes communications circuitry suitable for communicating with an application server or other networks or servers. The instructions for carrying out the above-mentioned functionality may be stored on the application server. Communications circuitry may include a cable modem, an integrated-services digital network (ISDN) modem, a digital subscriber line (DSL) modem, a telephone modem, an ethernet card, or a wireless modem for communications with other equipment, or any other suitable communications circuitry. Such communications may involve the Internet or any other suitable communications networks or paths. In addition, communications circuitry may include circuitry that enables peer-to-peer communication of user equipment devices, or communication of user equipment devices in locations remote from each other (described in more detail below).

[0032] Memory may be an electronic storage device such as storage 308 that is part of control circuitry 304. As referred to herein, the phrase "electronic storage device" or "storage device" should be understood to mean any device for storing electronic data, computer software, or firmware, such as random-access memory, read-only memory, hard drives, optical drives, solid state devices, quantum storage devices, gaming consoles, gaming media, any other suitable fixed or removable storage devices, and/or any combination of the same. Storage 308 may be used to store various types of content described herein as well as media guidance data described above. Nonvolatile memory may also be used (e.g., to launch a boot-up routine and other instructions). Cloud-based storage, for example, may be used to supplement storage 408 or instead of storage 308.

[0033] A user may send instructions to control circuitry 304 using user input interface 310. User input interface 310, display 312, or both may include a touchscreen configured to provide a display and receive haptic input. For example, the touchscreen may be configured to receive haptic input from a finger, a stylus, or both. In some embodiments, user device 300 may include a front-facing screen and a rear-facing screen, multiple front screens, or multiple angled screens. In some embodiments, user input interface 310 includes a remote-control device having one or more microphones, buttons, keypads, any other components configured to receive user input, or combinations thereof. For example, user input interface 310 may include a handheld remote-control device having an alphanumeric keypad and option buttons. In a further example, user input interface 310 may include a handheld remote-control device having a microphone and control circuitry configured to receive and identify voice commands and transmit information to set-top box 316.

[0034] Audio equipment 314 may be provided as integrated with other elements of each one of user device 300 and user equipment system 301 or may be stand-alone units. The audio component of videos and other content displayed on display 312 may be played through speakers of audio equipment 314. In some embodiments, the audio may be distributed to a receiver (not shown), which processes and outputs the audio via speakers of audio equipment 314. In some embodiments, for example, control circuitry 304 is configured to provide audio cues to a user, or other audio feedback to a user, using speakers of audio equipment 314. Audio equipment 314 may include a microphone configured to receive audio input such as voice commands and speech (e.g., including voice queries). For example, a user may speak letters or words that are received by the microphone and converted to text by control circuitry 304. In a further example, a user may voice commands that are received by the microphone and recognized by control circuitry 304.

[0035] An application (e.g., for managing voice queries) may be implemented using any suitable architecture. For example, a stand-alone application may be wholly implemented on each one of user device 300 and user equipment system 301. In some such embodiments, instructions for the application are stored locally (e.g., in storage 308), and data for use by the application is downloaded on a periodic basis (e.g., from an out-of-band feed, from an Internet resource, or using another suitable approach). Control circuitry 304 may retrieve instructions for the application from storage 308 and process the instructions to generate any of the displays discussed herein. Based on the

processed instructions, control circuitry 304 may determine what action to perform when input is received from input interface 310. For example, movement of a cursor on a display up/down may be indicated by the processed instructions when input interface 310 indicates that an up/down button was selected. An application and/or any instructions for performing any of the embodiments discussed herein may be encoded on computer-readable media. Computer-readable media includes any media capable of storing data. The computer-readable media may be transitory, including, but not limited to, propagating electrical or electromagnetic signals, or may be non-transitory including, but not limited to, volatile and non-volatile computer memory or storage devices such as a hard disk, floppy disk, USB drive, DVD, CD, media card, register memory, processor cache, Random Access Memory (RAM), etc.

[0036] In some embodiments, the application is a client/server-based application. Data for use by a thick or thin client implemented on each one of user device 300 and user equipment system 301 is retrieved on demand by issuing requests to a server remote from each one of user device 300 and user equipment system 301. For example, the remote server may store the instructions for the application in a storage device. The remote server may process the stored instructions using circuitry (e.g., control circuitry 304) and generate the displays discussed above and below. The client device may receive the displays generated by the remote server and may display the content of the displays locally on user device 300. This way, the processing of the instructions is performed remotely by the server while the resulting displays, which may include text, a keyboard, or other visuals, are provided locally on user device 300. User device 300 may receive inputs from the user via input interface 310 and transmit those inputs to the remote server for processing and generating the corresponding displays. For example, user device 300 may transmit a communication to the remote server indicating that an up/down button was selected via input interface 310. The remote server may process instructions in accordance with that input and generate a display of the application corresponding to the input (e.g., a display that moves a cursor up/down). The generated display is then transmitted to user device 300 for presentation to the user.

[0037] In some embodiments, the application is downloaded and interpreted or otherwise run by an interpreter or virtual machine (e.g., run by control circuitry 304). In some embodiments, the application may be encoded in the ETV Binary Interchange Format (EBIF), received by control circuitry 304 as part of a suitable feed, and interpreted by a user agent running on control circuitry 304. For example, the

application may be an EBIF application. In some embodiments, the application may be defined by a series of JAVA-based files that are received and run by a local virtual machine or other suitable middleware executed by control circuitry 304.

[0038] In some embodiments, user device 300, illustrated as a wireless-enabled device, may be coupled to a communications network (e.g., connected to the Internet). For example, user device 300 may be coupled to a communications network via a communications path (e.g., which may include an access point). In some embodiments, user equipment system 301 may be a computing device coupled to a communications network via a wireless or wired connection (e.g., a LAN, or any other suitable communications link). A communications network may be one or more networks including the Internet, a mobile phone network, mobile voice or data network (e.g., a 4G or LTE network), cable network, public switched telephone network, or other types of communications network or combinations of communications networks. Communications paths may include one or more communications paths, such as a satellite path, a fiber-optic path, a cable path, a path that supports Internet communications, free-space connections (e.g., for broadcast or other wireless signals), or any other suitable wired or wireless communications path or combination of such paths. User device 300, user equipment system 301, or both may communicate directly with each other via communications paths, such as those described above, as well as other short-range point-to-point communications paths, such as USB cables, IEEE 1394 cables, wireless paths (e.g., Bluetooth, infrared, IEEE 802-11x, etc.), or other short-range communication via wired or wireless paths. BLUETOOTH is a certification mark owned by Bluetooth SIG, INC. Devices may also communicate with each other directly through an indirect path via a communications network.

[0039] In some embodiments, communications between user device 300, user equipment system 301, or both and a network device (e.g., a server, an applications server, a database) may be exchanged over one or more communications paths. A plurality of network entities may exist and be in communication with user device 300, user equipment system 301, or both. In some embodiments, a network device implements an application that communicates with instances of applications at many user devices (e.g., user device 300 or user equipment system 301).

[0040] In some embodiments, the network device includes one or more types of stored information, including, for example, entity information, metadata, content, historical communications and search records, user preferences, user profile information,

advertisement information, any other suitable information, or any combination thereof that may be retrieved by user device 300, user equipment system 301, or both. The network device may include an applications-hosting database or server, plug-ins, a software developers kit (SDK), an applications programming interface (API), or other software tools configured to provide software (e.g., as downloaded to a user device), run software remotely (e.g., hosting applications accessed by user devices), or otherwise provide applications support to applications of user device 300. In some embodiments, information from a network device is provided to user device 300 or user equipment system 301 using a client/server approach.

[0041] In some embodiments, an application may be implemented on user device 300, user equipment system 301, or both. For example, the application may be implemented as software or a set of executable instructions, which may be stored in storage 308 of user device 300, user equipment system 301, or both and executed by control circuitry of the respective devices. For example, an application may be implemented partially as a client application on user device 300 and partially on a remote server as a server application running on control circuitry of the remote server (e.g., control circuitry of a network device). When executed by control circuitry of the remote server, the application may instruct the control circuitry to generate a display and transmit the generated display to user device 300. The server application may instruct the control circuitry of the remote device to transmit data for storage on user device 300. The client application may instruct control circuitry of the receiving user device to generate the application displays.

[0042] In some embodiments, user device 300, user equipment system 301, or both, are part of a cloud-based arrangement. The cloud provides access to services, such as information storage, advertising, shopping, searching, messaging, or social networking services, among other examples, as well as access to any content described above, for user devices. Services can be provided in the cloud through cloud-computing service providers, or through other providers of online services. For example, the cloud-based services can include a storage service, a sharing site, a social networking site, a search engine, or other services via which user-sourced content is distributed for viewing by others on connected devices. These cloud-based services may allow a user device to store information to the cloud and to receive information from the cloud rather than storing information locally and accessing locally stored information. Cloud resources may be accessed by a user device using, for example, a web browser, a messaging

application, a social media application, an authentication application, an application requiring authentication, a desktop application, a mobile application, any other suitable application, or any combination of applications.

[0043] FIG. 4 is a flowchart of an illustrative process 400 for generating a voice command recommendation, in accordance with some embodiments of the present disclosure. In some embodiments, as described below, process 400 may be performed by an application implemented on any suitable hardware. For example, an application may perform process 400, implemented on any suitable hardware such as interactive system 220 of FIG. 2, user device 300 of FIG. 3, user equipment system 301 of FIG. 3, any other suitable device, or any combination thereof.

[0044] Step 402 includes the application determining at least one predicted command expected to be received. The application may determine the at least one predicted command at a predetermined time, in response to an event, based on one or more conditions or triggers, based on user behavior, or a combination thereof. In an illustrative example, step 402 may be performed by interactive system 220 or command predictor 224 thereof.

[0045] In some embodiments, at step 402, the application determines the at least one predicted command based on historical information of the user's viewing behavior, command history, preferences, any other suitable information corresponding to the user, or any combination thereof. For example, the application may store or retrieve information such as a time history of commands (e.g., voice and non-voice commands, or lack thereof), command history relative to viewing behavior (e.g., time of a command relative to program start/stop time or other time marker), viewing history (e.g., which content is viewed and when), any other user information, or any combination thereof. In an illustrative example, at step 402 the application, as implemented by interactive system 220, may retrieve information from user information 290. In an illustrative example, at step 402, the application may retrieve historical use information from user information 290 that includes one or more received commands corresponding to the one or more events. In some embodiments, for example, the application may identify one or more time markers (e.g., stored in storage 308, user information 290, or both) and a corresponding action or command the user has initiated (e.g., and is likely to initiate again). For example, a database may include an entry having a time stamp of 9:00pm and a corresponding action of "switching channel to PBS," which has occurred a total of N times. In some embodiments, the time stamp is included as a time marker (e.g.,

included with information for populating a grid of content listings). When the application detects a clock time of 9:00pm (e.g., retrieved from clock/time information 280), the application may access the database to determine the corresponding command. For example, referencing display 100 of FIG. 1, the application may detect a time marker of 8:59 pm, and determine the command that corresponds to notification 101 (e.g., changing the channel). In some embodiments, for example, the application may identify one or more triggers based on user action (e.g., stored in storage 308, user information 290, or both), and a corresponding action or command the user has initiated. For example, a database may include an entry having a user action of "changing channels a predetermined number of times in a period" and a corresponding action of "navigate to an on-demand program listing," which has occurred a total of N times. In some embodiments, the action trigger is stored in memory (e.g., storage 308) accessible by the application. When the application detects that the user has changed the channel the predetermined number of times in the time period, the application may identify this trigger and access the database to determine the corresponding command. For example, the application may detect five channel changes in a minute and determine a corresponding command of "navigate to an on-demand program listing." In some embodiments, the application may monitor the time duration between voice commands, the time since the last voice command, a number or frequency of voice commands, or a combination thereof. The application may determine a voice command suggestion to motivate the user to use or explore voice commands. The voice command suggestion may be a common command, a relatively simple command, a popular command, or any other command.

[0046] In an illustrative example, the application may determine that a current time corresponds to a predetermined time marker at step 402. The application may then, at step 402, determine the least one predicted command by retrieving preference information corresponding to the at least one predicted command (e.g., from user information 290) and identifying the at least one predicted command as being preferred for voice input based on preference information. In a further illustrative example, the application may identify a time marker based on historical display activity, historical received inputs, or both.

[0047] In some embodiments, at step 402, the application determines the at least one predicted command based on statistical information of a plurality of users' viewing behavior, command histories, preferences, any other suitable information corresponding

to the plurality of users, or any combination thereof. The application may retrieve data from a database storing historical information (e.g., to extract a command), transmit data to a database storing historical information (e.g., to build a record of command-based activities), or both. For example, the application may store information for a plurality of users in a similar from as that described above for a single user. Further, the application may extract statistical information from the record of the plurality of users' interactions. To illustrate, the application may store information about which channels users primarily change to at certain times, and use the most popular (e.g., most frequent) command as a predicted command. In a further illustrative example, the application may determine that 60% of users increase the volume when changing to a music channel, and thus may predict the command "increase the volume" when the user changes the channel to a music channel (e.g., the channel change command being the trigger for the volume command). In an illustrative example, the application as implemented by interactive system 220, may retrieve information regarding a plurality of users from user information 290.

[0048] In some embodiments, at step 402, the application determines the at least one predicted command based on environmental factors, or other suitable circumstance information. In an illustrative example, at step 402, the application may identify one or more events corresponding to a viewing environment and determine the at least one predicted command expected to be received, based at least in part on the one or more events. The application may retrieve data from a database storing condition information (e.g., to extract a command), transmit data to a database storing condition information (e.g., to build a record of condition-based activities), or both. In an illustrative example, the application as implemented by interactive system 220 may retrieve information from condition information 270. The application may retrieve information (e.g., from condition information 270) based on one or more devices, sensors, or other suitable sources. For example, the application may be configured to communicate signals to and from a remote control or other interactive device, a peripheral device, or any other suitable device, to determine the device's position. The application may determine one or more users' positions, one or more devices' positions, the presence of one or more users, the presence of one or more devices, one or more events, one or more actions performed on a device, any other suitable condition information, or any combination thereof. For example, the application may determine other users in proximity to a first user. For example, if two users are in proximity to the display device, the application may identify conditions

wherein one user leaves and may generate a corresponding voice command recommendation (e.g., pause the program until they return to proximity, or increase the volume to reach the greater distance). In a further example, if one user is in proximity to the display device, the application may identify conditions wherein another user arrives and may generate a corresponding voice command recommendation (e.g., pause the program until they get situated after a predetermined time).

[0049] In an illustrative example, at step 402, the application may determine the at least one predicted command expected to be received by identifying one or more events, retrieving historical use information, and determining the at least one predicted command based on the one or more events and based on the historical use information. To illustrate, the application may identify a phone call ring, determine that the user normally mutes the volume of the television in response to phone calls, and determines a predicted command of "decrease volume."

[0050] Step 404 includes the application generating a voice command recommendation. The voice command recommendation corresponds to the at least one predicted command of step 402. In an illustrative example, the application as implemented by interactive system 220, may use recommendation generator 225 to perform 404. In an illustrative example, the application may generate a voice command recommendation at step 404 by causing a display device to modify a display to include a visual indicator corresponding to the voice command recommendation.

The application may generate a recommendation in the form of a notification, as illustrated by notifications 101 and 151 of FIG. 1. The recommendation of step 404 may include a description of the voice-command, introductory language (e.g., a greeting or other softening language), illustrative voice responses that are recognized, a time limit for responding, context information (e.g., the trigger or reason for the recommendation, the time since the last voice command, an encouragement to try using voice commands), feedback (e.g., whether to provide such notifications, when to provide such notifications), any other suitable information, or any combination thereof. In some embodiments, the application constructs a notification that may include text, images, video, or a combination thereof based on the predicted command of step 402, information about the user (e.g., from user information 290), information about a condition or circumstance (e.g., from condition information 270), information about a content listing (e.g., a program title or thumbnail image), information about one or more other users (e.g., how popular the command is), reasons for using the voice command (e.g., speaking a single word replaces several

keystrokes on a remote control), any other suitable information, or any combination thereof. In some embodiments, the application determines how to present the notification to the user. For example, the application may format a text box overlay, as illustrated by notifications 101 and 151 of FIG. 1, for display on a display device (e.g., using display generator 221). In a further illustrative example, the application may generate the voice command recommendation by causing an audio device (e.g., a speaker) to generate an audible indicator corresponding to the voice command recommendation. In a further example, the application may generate a voice query to the user by generating speech based on the notification and generating audio output on a speaker or other audio equipment (e.g., audio equipment 314). In some embodiments, the predicted command may correspond to a non-voice command or series of non-voice commands, for which there is no corresponding single voice command. In some such embodiments, the application may determine a simpler or otherwise more limited voice command that corresponds to the non-voice command to some suitable extent. For example, a predicted command may include a series of actions such as navigate to an on-demand screen, select a channel of on-demand content, and view descriptions of the content (e.g., based on a user's history of browsing content). The application may recommend a voice command only for navigating to the on-demand screen to facilitate a relatively more concise interaction with the user (e.g., especially if the user does not use voice commands frequently). In some embodiments, the application generates more than one voice command recommendation. For example, a notification may include one or more recommended voice commands to achieve one or more functions.

[0051] In an illustrative example, the predicted command of step 402 includes an action, and the voice command recommendation of step 404 includes a recommendation to the user to achieve that command. In some embodiments, the predicted command and the recommended voice command are the same. For example, in some such embodiments, steps 402 and 404 may be combined (e.g., command predictor 224 and recommendation generator 225 may be combined).

[0052] Step 406 includes the application activating an audio interface that is configured to receive a voice input. In some embodiments, the audio interface is always configured to receive audio input. In some embodiments, the audio interface is configured to receive audio input when a user provides an indication to a user input interface (e.g., by selecting a soft button on a touchscreen to begin audio recording). In some embodiments, the audio interface is configured to receive audio input in response to generating a voice command recommendation at step 404. In some embodiments, the

audio interface is configured to receive audio input and begins recording when speech or other suitable audio signals are detected.

[0053] Step 408 includes the application receiving the voice input at the audio interface. In some embodiments, an audio interface (e.g., audio equipment 314, user input interface 310, or a combination thereof) may include a microphone or other sensor that receives audio input and generates an electronic signal. In some embodiments, the audio input is received at an analog sensor, which provides an analog signal that is conditioned, sampled, and digitized to generate an audio file. In some embodiments, the audio file is stored in memory (e.g., storage 308). The query application may include any suitable conditioning software or hardware for converting audio input to a stored audio file. For example, the application may apply one or more filters (e.g., low-pass, high-pass, notch filters, or band-pass filters), amplifiers, decimators, or other conditionings to generate the audio file. In a further example, the application may apply any suitable processing to a conditioned signal to generate an audio file such as compression, transformation (e.g., spectral transformation, wavelet transformation), normalization, equalization, truncation (e.g., in a time or spectral domain), any other suitable processing, or any combination thereof. In some embodiments, at step 408, the control circuitry may receive a voice command as an audio file stored in storage (e.g., storage 308), for further processing.

[0054] In some embodiments, the application may store snippets (i.e., clips of short duration) of received audio, and process the snippets. In some embodiments, the application stores relatively large segments of speech (e.g., more than 10 seconds) as an audio file, and processes the file. In some embodiments, the query application may process speech to detect words by using a continuous computation. For example, a wavelet transform may be performed on speech in real time, providing a continuous, if slightly time-lagged, computation of speech patterns (e.g., which could be compared to a reference to identify words). In some embodiments, the application may detect words, as well as which user uttered the words (e.g., voice recognition), in accordance with the present disclosure.

[0055] The application may receive voice input and process words at step 408 using any suitable word detection technique, speech detection technique, pattern recognition technique, signal processing technique, or any combination thereof. For example, the application may compare a series of signal templates to a portion of an audio signal to find whether a match exists (e.g., whether a particular word is included in the audio

signal). In a further example, the application may apply a learning technique to better recognize words in voice commands. For example, the application may gather feedback from a user on a plurality of voice commands, and accordingly use past data as a training set for making recommendations of voice commands.

[0056] Step 410 includes the application performing the at least one predicted command. The application performs the at least one predicted command of step 402, or a portion thereof, in response to receiving the voice input. In some embodiments, the application may wait a predetermined time for a response (e.g., 30 seconds or any other suitable time), and then cease displaying the notification. For example, if no response (e.g., voice input) is detected or otherwise recognizable (e.g., the speech detection is not successful in identifying a recognized response), the application may cease presenting the notification, modify the notification with an indicator that the time limit is impending, modify the notification to increase visibility (e.g., change color, size, or contrast of the notification to receive attention from the user), generate an audio indicator (e.g., generate a beep, sound, or speech to alert the user), or any combination thereof. In some embodiments, the application does not perform the predicted command because no voice input was detected or recognized, a negative voice input was received (e.g., "no" or "exit"), or a condition or trigger has changed (e.g., a modified notification may be generated). The application may perform the at least one predicted command immediately in response to receiving suitable voice input, at a predetermined time (e.g., change the channel at 9:00pm as illustrated by display 100 of FIG. 1), in response to an event occurrence, or at any other suitable time.

[0057] In an illustrative example, the application may determine a predicted command that includes changing the display from content corresponding to a first channel to content corresponding to a second channel at step 402, and may perform the at least one predicted command by displaying the content corresponding to the second channel at step 410.

[0058] In a further illustrative example, the application may determine a predicted command that includes changing a volume of an audio track associated with the display, and may perform the at least one predicted command by increasing or decreasing the volume of the audio track at step 410.

[0059] FIG. 5 shows a flowchart of illustrative process 500 for managing voice command interactions, in accordance with some embodiments of the present disclosure. In some embodiments, as described below, process 400 may be implemented on any

suitable hardware such as interactive system 220 of FIG. 2, user device 300 of FIG. 3, user equipment system 301 of FIG. 3, any other suitable device, or any combination thereof. In an illustrative example, the system may implement an application to perform process 500 (e.g., similar to the application described in the context of FIG. 4).

[0060] Step 502 includes the system monitoring interactions, conditions, time, any other suitable information, or any combination thereof to identify triggers. For example, the system may monitor keystrokes, non-voice commands, voice commands, a current time, a time duration, events, activity on one or more devices, any other suitable activity or information, or any combination thereof. For example, the system may retrieve potential triggers or criteria from information 590, transmit usage data to information 590, or both. Information 590 may include, for example, user information 290, condition information 270, clock/time information 280, any other suitable information, or any combination thereof.

[0061] Step 504 includes the system determining whether a trigger has been detected. In some embodiments, the system determines whether one or more triggers has occurred, and in response to detecting a trigger, the system may proceed to step 506 to determine a command. The system may monitor for triggers continuously, at intervals, in response to receiving input, or at any other suitable times.

[0062] Step 506 includes the system determining a command based on the detected trigger of step 502. In some embodiments, the system retrieves a command from information 590 based on the detected trigger at step 504. The command may correspond to the trigger based on any suitable criteria including, for example, a relationship in a database.

[0063] Step 508 includes the system generating a notification based on the command of step 506. The system may determine a recommended voice command and generate a visual or audio notification to the user (e.g., using a display device or audio device), at step 508.

[0064] Step 510 includes the system activating an interface (e.g., an audio interface). The system may perform step 510 before, after, or at the same time as step 508. For example, the system may perform step 510 in response to a trigger being detected at step 504, retrieval of information from information 590, generation of the notification at step 508, or at any other suitable time.

[0065] Step 512 includes the system determining whether a response has been detected at the interface activated at step 510. The system may monitor the interface for a

predetermined period of time, until a response is detected, until an event occurs (e.g., any command is received, voice or non-voice), any other suitable duration, or any combination thereof. For example, the system may process signals received at the interface using a speech detection technique. When the system identifies a recognizable voice input, the system may proceed to step 514 to response to the command.

[0066] Step 514 includes the system performing the command corresponding to the notification of step 508. The system may perform any suitable function at step 514 including changing a channel, changing or (un)muting a volume, turning a device on or off, pausing or playing a program, rewinding or forwarding a program, changing a display or display setting, any other suitable function, or any combination thereof.

[0067] Step 516 includes the system determining whether to modify, update, or replace the notification. For example, if at step 512, no response is detected, the system may determine whether to modify the notification at step 516. For example, the system may determine not to modify the system and return to monitoring interactions at step 502. In a further example, the system may determine to modify, update, or replace the notification of step 508 to urge the user to interact with the system, remind the user of the notification, or otherwise indicate the lack of recognizable and actionable voice response.

[0068] Step 518 includes the system storing suitable information based on interactions during or in the context of process 500. For example, a record of the performed command, lack of voice response, trigger detected, time information associated with any aspect of process 500, any other suitable information, or any combination thereof, may be stored in information 590 for inclusion in historical information (e.g., for later determination of commands and usage behavior). Accordingly, the system can construct a record of the user's behavior and preferences for future command predictions and voice command recommendations.

[0069] The above-described embodiments of the present disclosure are presented for purposes of illustration and not of limitation, and the present disclosure is limited only by the claims which follow. Furthermore, it should be noted that the features and limitations described in any one embodiment may be applied to any other embodiment herein, and flowcharts or examples relating to one embodiment may be combined with any other embodiment in a suitable manner, done in different orders, or done in parallel. In addition, the systems and methods described herein may be performed in real time. It should also be noted, the systems and/or methods described above may be applied to, or used in accordance with, other systems and/or methods.

This specification discloses embodiments which include, but are not limited to, the following:

1. A method for providing a voice command recommendation, the method comprising:

determining at least one predicted command expected to be received;

generating, using control circuitry, a voice command recommendation, wherein the voice command recommendation corresponds to the at least one predicted command;

activating an audio interface configured to receive a voice input;

receiving the voice input at the audio interface; and

performing the at least one predicted command in response to receiving the voice input.

2. The method of item 1, wherein:

the at least one predicted command comprises changing the display from content corresponding to a first channel to content corresponding to a second channel; and

performing the at least one predicted command comprises displaying the content corresponding to the second channel.

3. The method of item 1, wherein:

the at least one predicted command comprises changing a volume of an audio track associated with the display; and

performing the at least one predicted command comprises increasing or decreasing the volume of the audio track.

4. The method of item 1, further comprising determining that a current time corresponds to a predetermined time marker, wherein determining the least one predicted command comprises:

retrieving preference information corresponding to the at least one predicted command, and

identifying the at least one predicted command as being preferred for voice input based on preference information.

5. The method of item 4, further comprising identifying the time marker based on at least one of the group comprising historical display activity, historical received inputs, and both.

6. The method of item 1, wherein generating the voice command recommendation comprises modifying a display, generated on a display device, to include a visual indicator corresponding to the voice command recommendation.

7. The method of item 1, wherein generating the voice command recommendation comprises generating, using an audio device, an audible indicator corresponding to the voice command recommendation.

8. The method of item 1, wherein determining the at least one predicted command expected to be received comprises:

identifying one or more events;

retrieving historical use information; and

determining the at least one predicted command based on the one or more events and based on the historical use information.

9. The method of item 8, wherein the historical use information comprises one or more received commands corresponding to the one or more events.

10. The method of item 1, further comprising:

identifying one or more events corresponding to a viewing environment; and

determining the at least one predicted command expected to be received, based at least in part on the one or more events.

11. A system for providing a voice command recommendation, the system comprising:

an audio interface configured to receive audio input; and

control circuitry configured to:

determine at least one predicted command expected to be received;

generate a voice command recommendation, wherein the voice command recommendation corresponds to the at least one predicted command;

activate the audio interface;

receive a voice input from the audio interface; and

perform the at least one predicted command in response to receiving the voice input.

12. The system of item 11, wherein the at least one predicted command comprises changing the display from content corresponding to a first channel to content corresponding to a second channel, and the control circuitry is further configured to perform the at least one predicted command by displaying the content corresponding to the second channel.

13. The system of item 11, wherein the at least one predicted command comprises changing a volume of an audio track associated with the display, and the control circuitry is further configured to perform the at least one predicted command by increasing or decreasing the volume of the audio track.

14. The system of item 11, wherein the control circuitry is further configured to:

determine that a current time corresponds to a predetermined time marker; and

determine the least one predicted command by:

retrieving preference information corresponding to the at least one predicted command, and

identifying the at least one predicted command as being preferred for voice input based on preference information.

15. The system of item 14, wherein the control circuitry is further configured to identify the time marker based on at least one of the group comprising historical display activity, historical received inputs, and both.

16. The system of item 11, further comprising a display device, wherein the control circuitry is further configured to generate the voice command recommendation by causing the display device to modify a display to include a visual indicator corresponding to the voice command recommendation.

17. The system of item 11, further comprising an audio device, wherein the control circuitry is further configured to generate the voice command recommendation by causing the audio device to generate an audible indicator corresponding to the voice command recommendation.

18. The system of item 11, wherein the control circuitry is further configured to determine the at least one predicted command expected to be received by:

identifying one or more events;

retrieving historical use information; and

determining the at least one predicted command based on the one or more events and based on the historical use information.

19. The system of item 18, wherein the historical use information comprises one or more received commands corresponding to the one or more events.

20. The system of item 11, wherein the control circuitry is further configured to:

identify one or more events corresponding to a viewing environment; and

determine the at least one predicted command expected to be received, based at least in part on the one or more events.

21. A non-transitory computer-readable medium having instructions encoded thereon that when executed by control circuitry cause the control circuitry to:

determine at least one predicted command expected to be received;

generate a voice command recommendation, wherein the voice command recommendation corresponds to the at least one predicted command;

activate an audio interface configured to receive a voice input;

receive the voice input at the audio interface; and

perform the at least one predicted command in response to receiving the voice input.

22. The non-transitory computer-readable medium of item 21, wherein the at least one predicted command comprises changing the display from content corresponding to a first channel to content corresponding to a second channel,

further comprising instructions encoded thereon that when executed by control circuitry cause the control circuitry to perform the at least one predicted command by displaying the content corresponding to the second channel.

23. The non-transitory computer-readable medium of item 21, wherein the at least one predicted command comprises changing a volume of an audio track associated with the display, further comprising instructions encoded thereon that when executed by control circuitry cause the control circuitry to perform the at least one predicted command by increasing or decreasing the volume of the audio track.

24. The non-transitory computer-readable medium of item 21, further comprising instructions encoded thereon that when executed by control circuitry cause the control circuitry to:

determine that a current time corresponds to a predetermined time marker, and

determine the least one predicted command by:

retrieving preference information corresponding to the at least one predicted command, and

identifying the at least one predicted command as being preferred for voice input based on preference information.

25. The non-transitory computer-readable medium of item 24, further comprising instructions encoded thereon that when executed by control circuitry cause the control circuitry to identify the time marker based on at least one of the group comprising historical display activity, historical received inputs, and both.

26. The non-transitory computer-readable medium of item 21, further comprising instructions encoded thereon that when executed by control circuitry cause the control circuitry to generate the voice command recommendation by modifying a display, generated on a display device, to include a visual indicator corresponding to the voice command recommendation.

27. The non-transitory computer-readable medium of item 21, further comprising instructions encoded thereon that when executed by control circuitry cause the control circuitry to generate the voice command recommendation by generating,

using an audio device, an audible indicator corresponding to the voice command recommendation.

28. The non-transitory computer-readable medium of item 21, further comprising instructions encoded thereon that when executed by control circuitry cause the control circuitry to determine the at least one predicted command expected to be received by:

identifying one or more events;

retrieving historical use information; and

determining the at least one predicted command based on the one or more events and based on the historical use information.

29. The non-transitory computer-readable medium of item 28, wherein the historical use information comprises one or more received commands corresponding to the one or more events.

30. The non-transitory computer-readable medium of item 21, further comprising instructions encoded thereon that when executed by control circuitry cause the control circuitry to:

identify one or more events corresponding to a viewing environment; and

determine the at least one predicted command expected to be received, based at least in part on the one or more events.

31. A system for providing a voice command recommendation, the system comprising:

means for determining at least one predicted command expected to be received;

means for generating a voice command recommendation, wherein the voice command recommendation corresponds to the at least one predicted command;

means for activating an audio interface configured to receive a voice input;

means for receiving the voice input at the audio interface; and

means for performing the at least one predicted command in response to receiving the voice input.

32. The system of item 31, wherein:

the at least one predicted command comprises changing the display from content corresponding to a first channel to content corresponding to a second channel; and

the means for performing the at least one predicted command comprises means for displaying the content corresponding to the second channel.

33. The system of item 31, wherein:

the at least one predicted command comprises changing a volume of an audio track associated with the display; and

the means for performing the at least one predicted command comprises means for increasing or decreasing the volume of the audio track.

34. The system of item 31, further comprising means for determining that a current time corresponds to a predetermined time marker, wherein the means for determining the least one predicted command comprises:

means for retrieving preference information corresponding to the at least one predicted command; and

means for identifying the at least one predicted command as being preferred for voice input based on preference information.

35. The system of item 34, further comprising means for identifying the time marker based on at least one of the group comprising historical display activity, historical received inputs, and both.

36. The system of item 31, wherein the means for generating the voice command recommendation comprises means for modifying a display, generated on a display device, to include a visual indicator corresponding to the voice command recommendation.

37. The system of item 31, wherein the means for generating the voice command recommendation comprises means for generating, using an audio device, an audible indicator corresponding to the voice command recommendation.

38. The system of item 31, wherein the means for determining the at least one predicted command expected to be received comprises:

means for identifying one or more events;

means for retrieving historical use information; and

means for determining the at least one predicted command based on the one or more events and based on the historical use information.

39. The system of item 38, wherein the historical use information comprises one or more received commands corresponding to the one or more events.

40. The system of item 31, further comprising:

means for identifying one or more events corresponding to a viewing environment; and

means for determining the at least one predicted command expected to be received, based at least in part on the one or more events.

41. A method for providing a voice command recommendation, the method comprising:

determining at least one predicted command expected to be received;

generating, using control circuitry, a voice command recommendation, wherein the voice command recommendation corresponds to the at least one predicted command;

activating an audio interface configured to receive a voice input;

receiving the voice input at the audio interface; and

performing the at least one predicted command in response to receiving the voice input.

42. The method of item 41, wherein:

the predicted command comprises changing the display from content corresponding to a first channel to content corresponding to a second channel; and

performing the at least one predicted command comprises displaying the content corresponding to the second channel.

43. The method of any of items 41 and 42, wherein:

the predicted command comprises changing a volume of an audio track associated with the display; and

performing the at least one predicted command comprises increasing or decreasing the volume of the audio track.

44. The method of any of items 41-43, further comprising determining that a current time corresponds to a predetermined time marker, wherein determining the least one predicted command comprises:

retrieving preference information corresponding to the at least one predicted command; and

identifying the at least one predicted command as being preferred for voice input based on preference information.

45. The method of item 44, further comprising identifying the time marker based on at least one of the group comprising historical display activity, historical received inputs, and both.

46. The method of any of items 41-45, wherein generating the voice command recommendation comprises modifying a display, generated on a display device, to include a visual indicator corresponding to the voice command recommendation.

47. The method of any of items 41-46, wherein generating the voice command recommendation comprises generating, using an audio device, an audible indicator corresponding to the voice command recommendation.

48. The method of any of items 41-47, wherein determining the at least one predicted command expected to be received comprises:

identifying one or more events;

retrieving historical use information; and

determining the at least one predicted command based on the one or more events and based on the historical use information.

49. The method of item 48, wherein the historical use information comprises one or more received commands corresponding to the one or more events.

50. The method of any of items 41-49, further comprising:

identifying one or more events corresponding to a viewing environment; and

determining the at least one predicted command expected to be received, based at least in part on the one or more events.

What is Claimed is:

1. A method for providing a voice command recommendation, the method comprising:

determining at least one predicted command expected to be received;

generating, using control circuitry, a voice command recommendation, wherein the voice command recommendation corresponds to the at least one predicted command;

activating an audio interface configured to receive a voice input;

receiving the voice input at the audio interface; and

performing the at least one predicted command in response to receiving the voice input.

2. The method of claim 1, wherein:

the predicted command comprises changing the display from content corresponding to a first channel to content corresponding to a second channel; and

performing the at least one predicted command comprises displaying the content corresponding to the second channel.

3. The method of any of claims 1 and 2, wherein:

the predicted command comprises changing a volume of an audio track associated with the display; and

performing the at least one predicted command comprises increasing or decreasing the volume of the audio track.

4. The method of any of claims 1-3, further comprising determining that a current time corresponds to a predetermined time marker, wherein determining the least one predicted command comprises:

retrieving preference information corresponding to the at least one predicted command, and

identifying the at least one predicted command as being preferred for voice input based on preference information.

5. The method of claim 4, further comprising identifying the time marker based on at least one of the group comprising historical display activity, historical received inputs, and both.

6. The method of any of claims 1-5, wherein generating the voice command recommendation comprises modifying a display, generated on a display device, to include a visual indicator corresponding to the voice command recommendation.

7. The method of any of claims 1-6, wherein generating the voice command recommendation comprises generating, using an audio device, an audible indicator corresponding to the voice command recommendation.

8. The method of any of claims 1-7, wherein determining the at least one predicted command expected to be received comprises:

identifying one or more events;

retrieving historical use information; and

determining the at least one predicted command based on the one or more events and based on the historical use information.

9. The method of claim 8, wherein the historical use information comprises one or more received commands corresponding to the one or more events.

10. The method of any of claims 1-9, further comprising:

identifying one or more events corresponding to a viewing environment; and

determining the at least one predicted command expected to be received, based at least in part on the one or more events.

11. A system for providing a voice command recommendation, the system comprising:

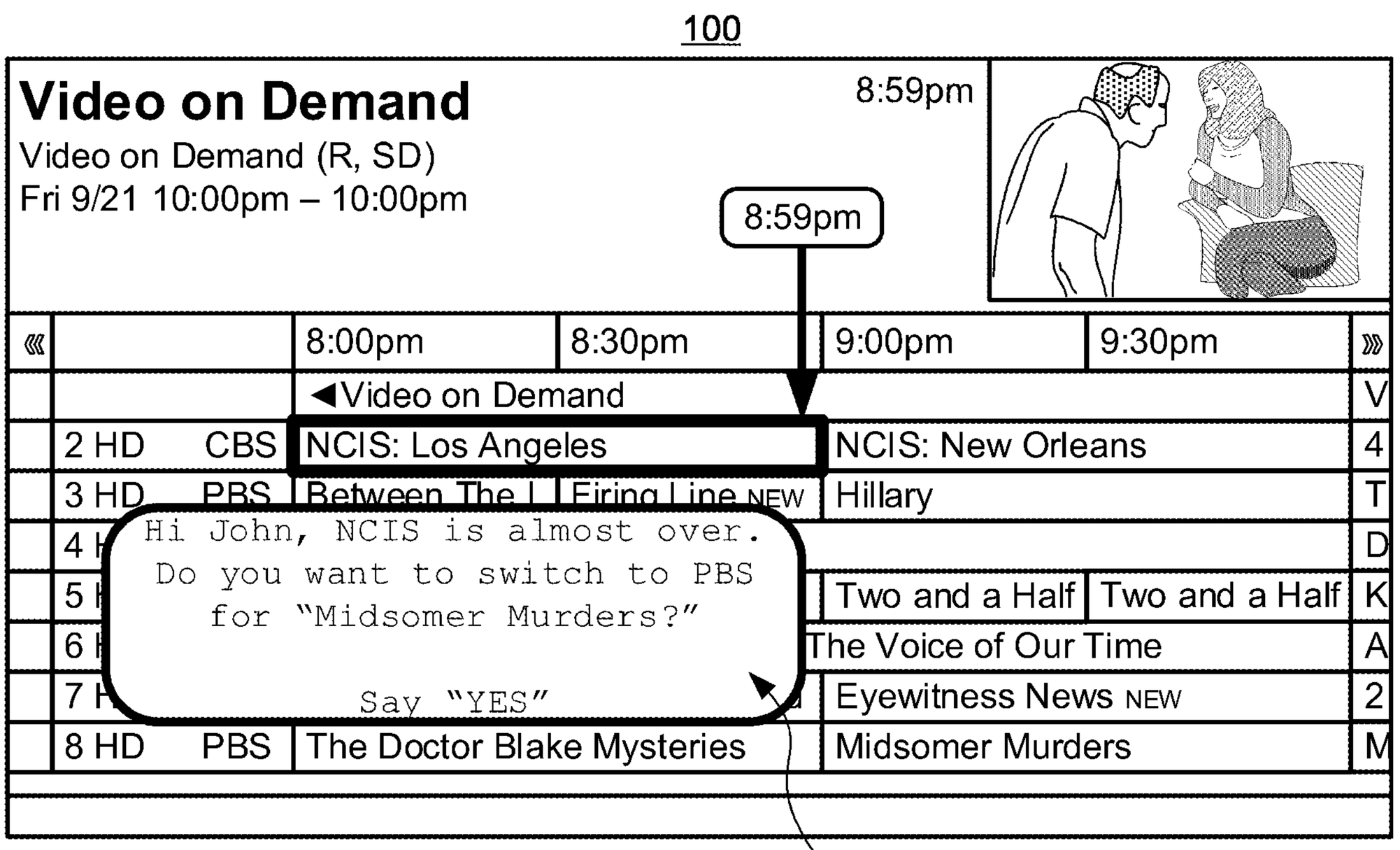
memory; and

means for implementing the steps of the method of any of claims 1 to 10.

12. A non-transitory computer-readable medium having instructions encoded thereon that when executed by control circuitry enable the control circuitry to execute the steps of the method of any of claims 1 to 10.

13. A system for providing a voice command recommendation, the system comprising:

means for implementing the steps of the method of any of claims 1 to 10.


100
Video on Demand
Video on Demand (R, SD)
Fri 9/21 10:00pm – 10:00pm
8:59pm
8:59pm
« 8:00pm 8:30pm 9:00pm 9:30pm »
◀Video on Demand V
2 HD CBS NCIS: Los Angeles NCIS: New Orleans 4
3 HD PBS Between The L Firing Line NEW Hillary T
4 H D
5 H Two and a Half Two and a Half K
6 H The Voice of Our Time A
7 H Eyewitness News NEW 2
8 HD PBS The Doctor Blake Mysteries Midsomer Murders M
Hi John, NCIS is almost over.
Do you want to switch to PBS
for "Midsomer Murders?"
Say "YES"
101

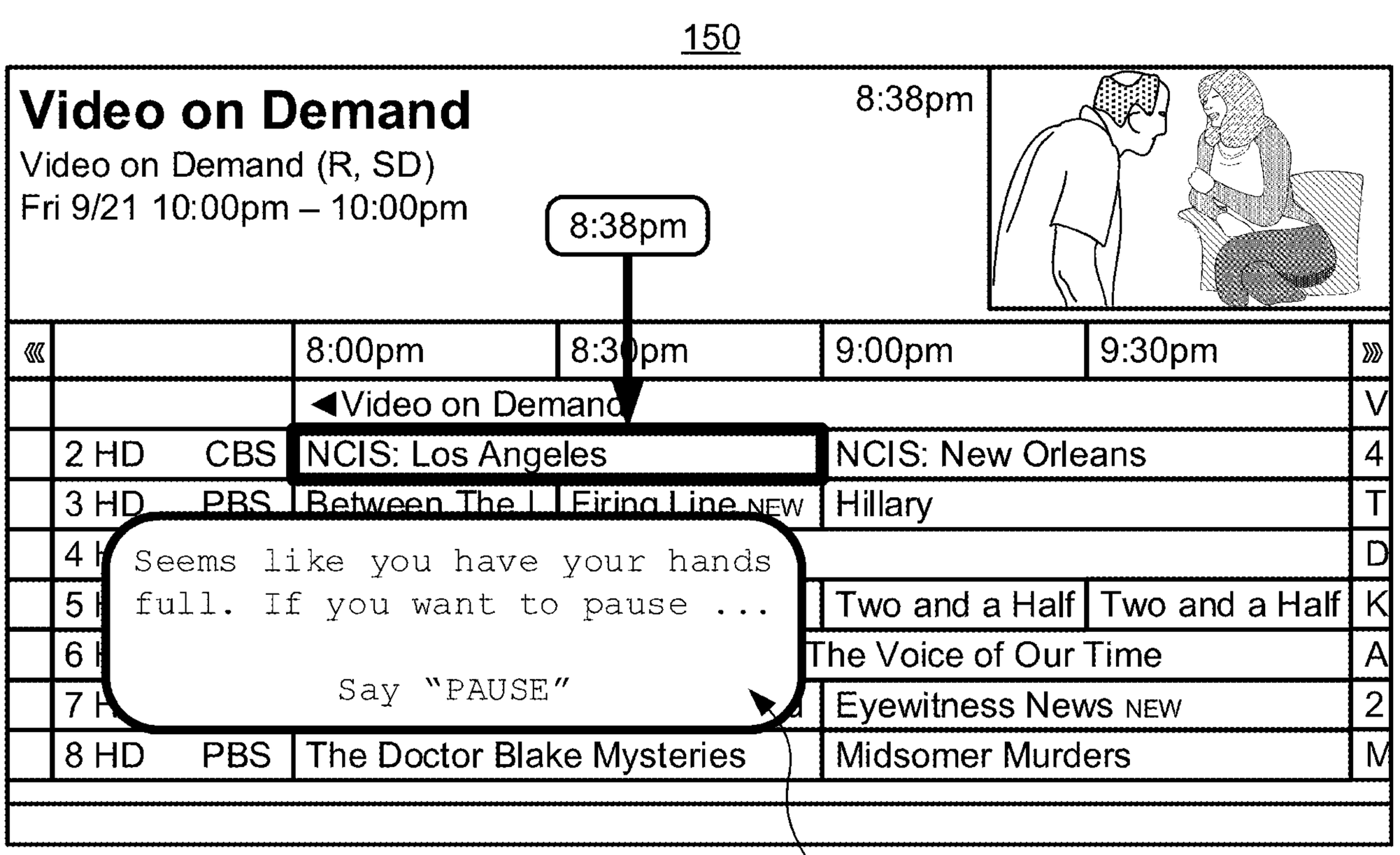

150
Video on Demand
Video on Demand (R, SD)
Fri 9/21 10:00pm – 10:00pm
8:38pm
8:38pm
« 8:00pm 8:30pm 9:00pm 9:30pm »
◀Video on Demand V
2 HD CBS NCIS: Los Angeles NCIS: New Orleans 4
3 HD PBS Between The L Firing Line NEW Hillary T
4 H D
5 H Two and a Half Two and a Half K
6 H The Voice of Our Time A
7 H Eyewitness News NEW 2
8 HD PBS The Doctor Blake Mysteries Midsomer Murders M
Seems like you have your hands
full. If you want to pause ...
Say "PAUSE"
151

FIG. 1

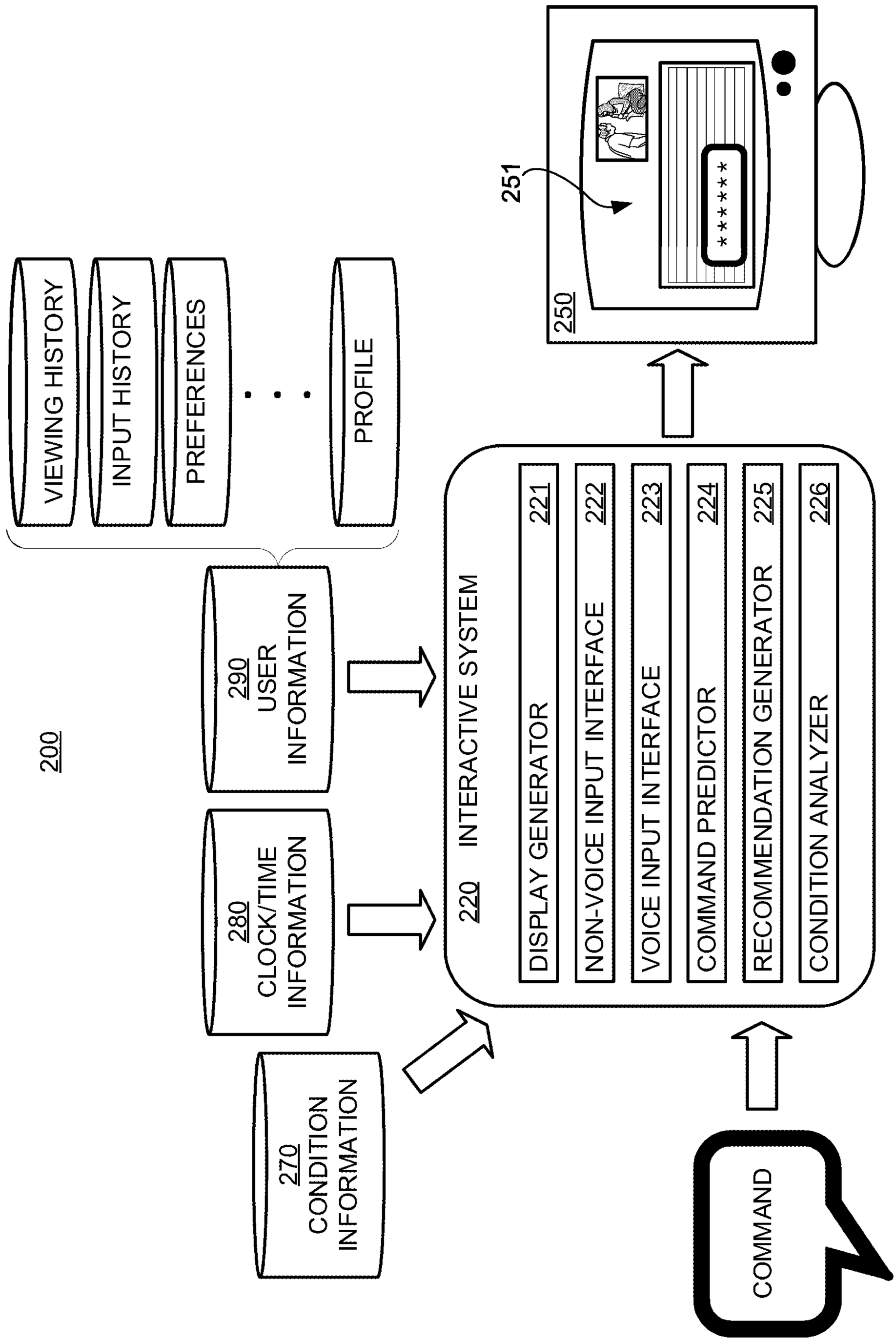

200
COMMAND
DISPLAY
220 INTERACTIVE SYSTEM
DISPLAY GENERATOR
NON-VOICE INPUT INTERFACE
VOICE INPUT INTERFACE
COMMAND PREDICTOR
RECOMMENDATION GENERATOR
CONDITION ANALYZER
221
222
223
224
225
226
280 CLOCK/TIME INFORMATION
290 USER INFORMATION
270 CONDITION INFORMATION
VIEWING HISTORY
INPUT HISTORY
PREFERENCES
PROFILE
250
251

FIG. 2

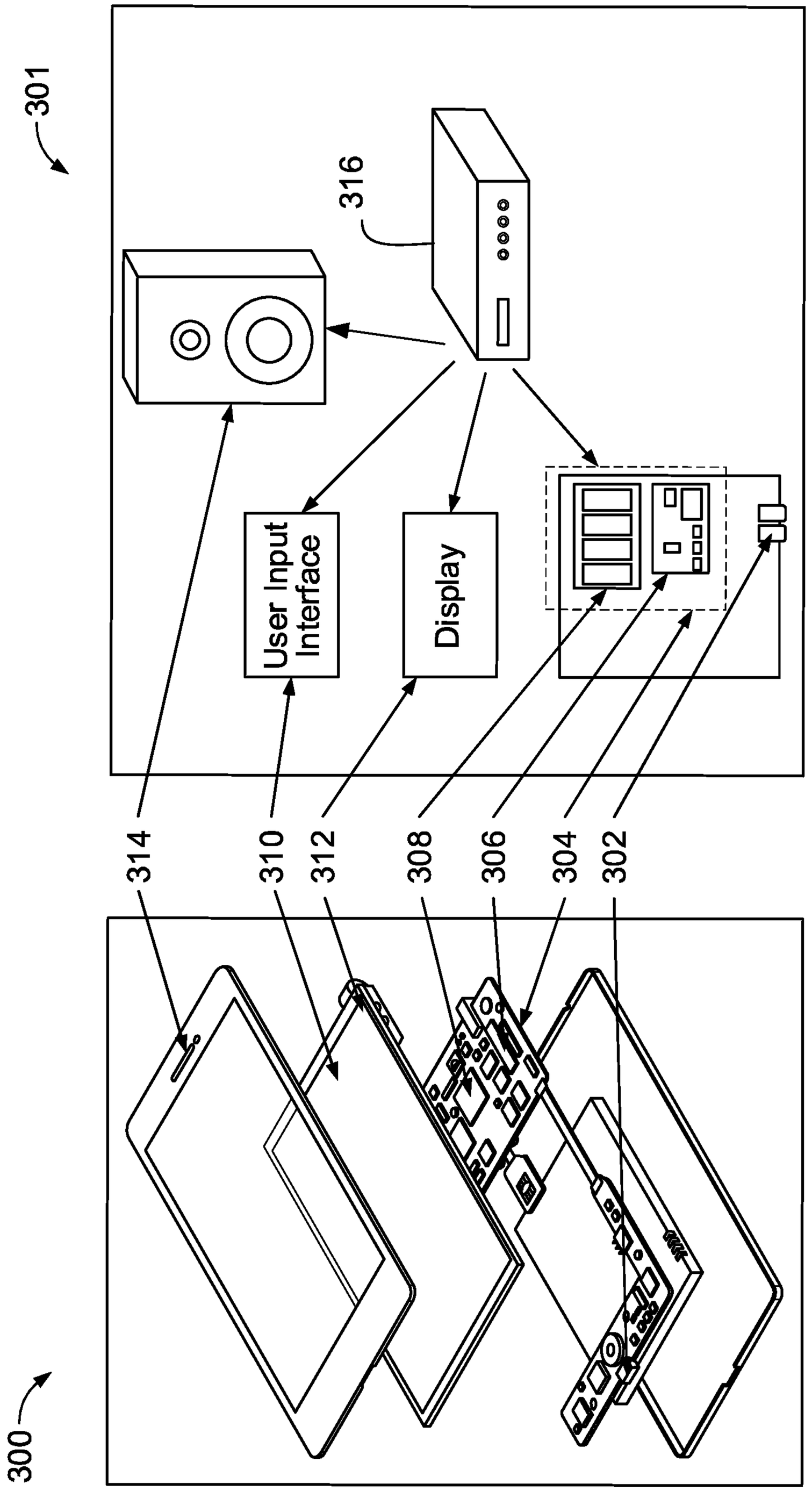

301
316
User Input Interface
Display
314
310
312
308
306
304
302
300

FIG. 3

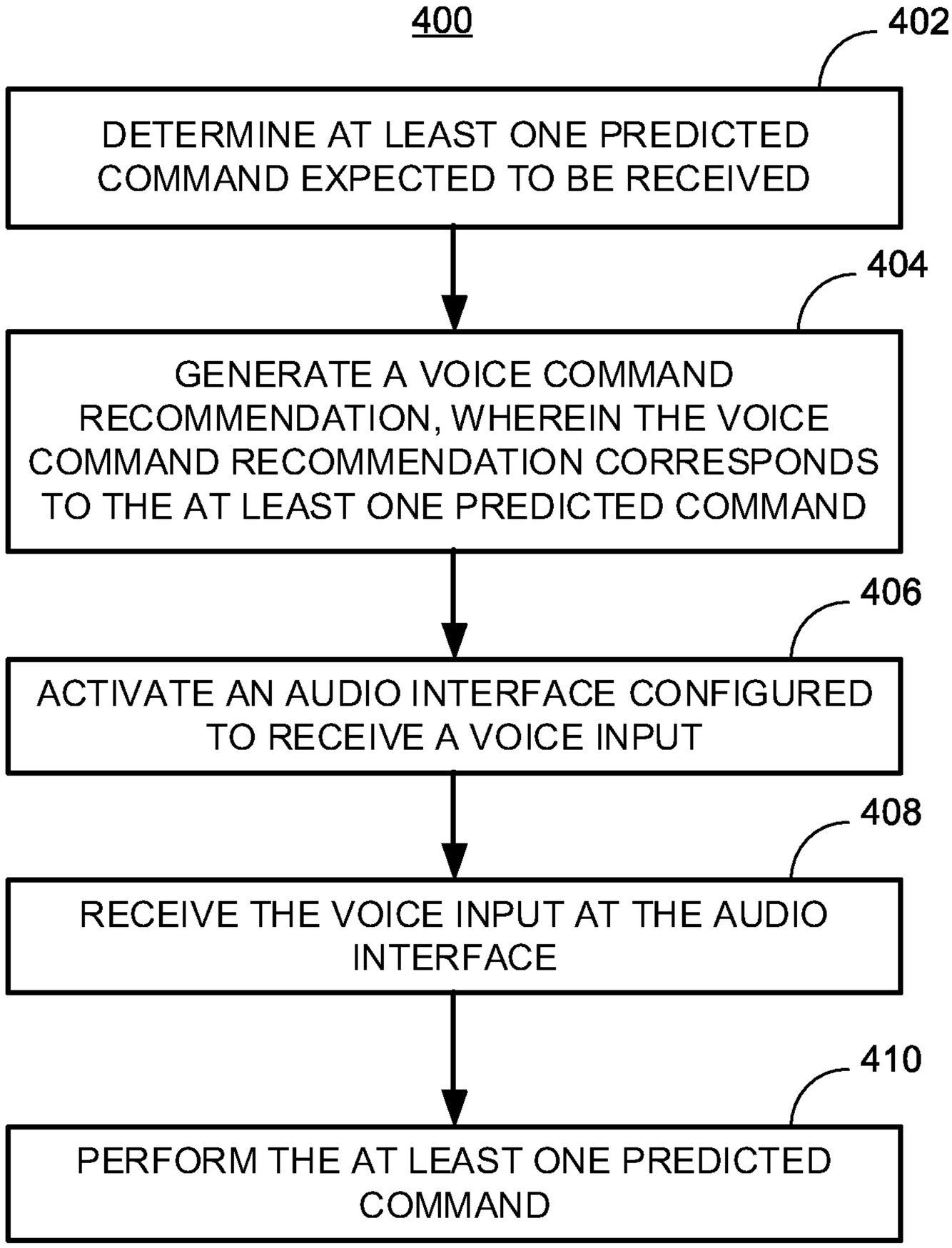

400
402
DETERMINE AT LEAST ONE PREDICTED COMMAND EXPECTED TO BE RECEIVED
404
GENERATE A VOICE COMMAND RECOMMENDATION, WHEREIN THE VOICE COMMAND RECOMMENDATION CORRESPONDS TO THE AT LEAST ONE PREDICTED COMMAND
406
ACTIVATE AN AUDIO INTERFACE CONFIGURED TO RECEIVE A VOICE INPUT
408
RECEIVE THE VOICE INPUT AT THE AUDIO INTERFACE
410
PERFORM THE AT LEAST ONE PREDICTED COMMAND

FIG. 4

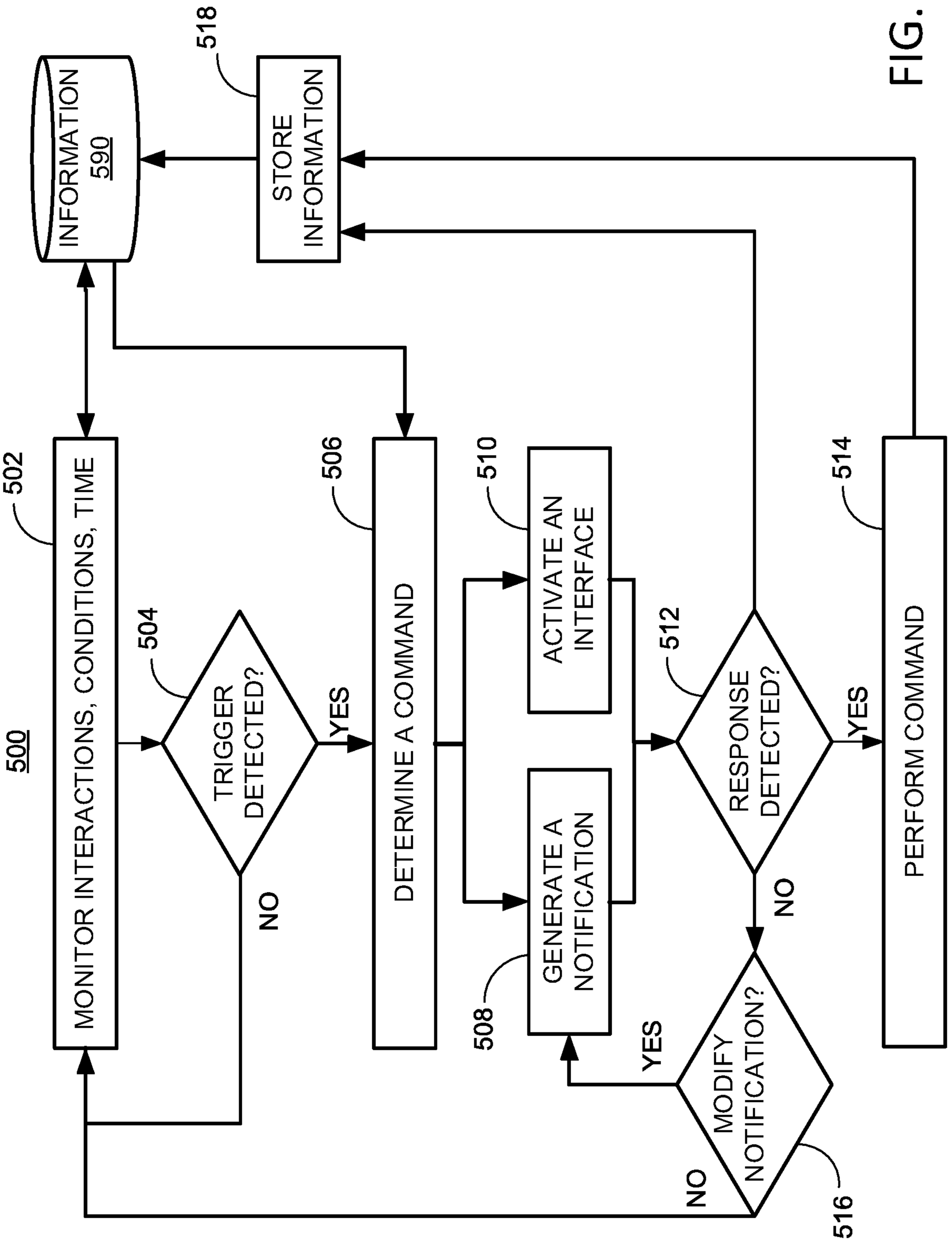

500
MONITOR INTERACTIONS, CONDITIONS, TIME
502
TRIGGER DETECTED?
504
NO
YES
DETERMINE A COMMAND
506
GENERATE A NOTIFICATION
508
ACTIVATE AN INTERFACE
510
RESPONSE DETECTED?
512
NO
YES
MODIFY NOTIFICATION?
516
NO
YES
PERFORM COMMAND
514
STORE INFORMATION
518
INFORMATION
590

FIG. 5

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2020/064048

A. CLASSIFICATION OF SUBJECT MATTER
INV. G10L15/22
ADD.

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G10L

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

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2019/066677 A1 (JAYGARL HO JUN [KR] ET AL) 28 February 2019 (2019-02-28) abstract; figures 1-26 paragraphs [0167] - [0199] paragraphs [0228] - [0233] paragraphs [0241] - [0297] -----	1-13
X	US 2018/308480 A1 (JANG DONG-HO [KR] ET AL) 25 October 2018 (2018-10-25) abstract; figures 1-29 paragraphs [0397] - [0452] -----	1-13

-/--

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
25 March 2021	07/04/2021

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Képesi, Marián

2

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2020/064048

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	Mcroberts Sarah ET AL: "Exploring Interactions with Voice-Controlled TV", , 14 May 2019 (2019-05-14), pages 1-11, XP055790070, Retrieved from the Internet: URL:https://arxiv.org/pdf/1905.05851.pdf [retrieved on 2021-03-25] abstract; figures 1, 2 pages 1, 2, 3 pages 7, 8 -----	1-13

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2020/064048

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2019066677 A1	28-02-2019	CN 110998720 A	10-04-2020
		EP 3652734 A1	20-05-2020
		KR 20190021143 A	05-03-2019
		US 2019066677 A1	28-02-2019
		WO 2019039834 A1	28-02-2019
US 2018308480 A1	25-10-2018	EP 3612925 A1	26-02-2020
		KR 20180117485 A	29-10-2018
		US 2018308480 A1	25-10-2018
		WO 2018194268 A1	25-10-2018