September 29, 2017

USER CONTROL OF CONTENT PLAYBACK BY AN INTERACTIVE ASSISTANT

Google Inc.

Follow this and additional works at: http://www.tdcommons.org/dpubs_series

Recommended Citation
Inc., Google, "USER CONTROL OF CONTENT PLAYBACK BY AN INTERACTIVE ASSISTANT", Technical Disclosure Commons, (September 29, 2017)
http://www.tdcommons.org/dpubs_series/697

This work is licensed under a Creative Commons Attribution 4.0 License.
This Article is brought to you for free and open access by Technical Disclosure Commons. It has been accepted for inclusion in Defensive Publications Series by an authorized administrator of Technical Disclosure Commons.
USER CONTROL OF CONTENT PLAYBACK
BY AN INTERACTIVE ASSISTANT

ABSTRACT

An interactive assistant, referred to herein as “an interactive assistant,” “a virtual assistant,” or simply “an assistant,” may be configured to dynamically adjust the playback of content (e.g., audible content output by the interactive assistant) in response to receiving a user command. Over the course of time, a user may engage in various different interactions with the assistant by submitting queries, performing searches, requesting select information, providing instructions, and the like. At any point in time, during playback of content provided by the interactive assistant, the user may issue a command to adjust the playback of such content. For example, the user may issue an audible command to instruct the interactive assistant to pause the playback of content, and may later issue a subsequent command to resume or continue playback of the content from the point at which playback was previously paused. As another example, the user may issue a command that causes the interactive assistant to skip over certain playback content or to repeat certain content that was previously output. In such fashion, and based on user feedback, the interactive assistant is capable of dynamically navigating to various points in a timeline of content that is played back to the user by the interactive assistant.

DESCRIPTION

An interactive assistant, such as shown in the example of Figure 1 below, may be included in a computing system that is configured to interact with one or more users. The computing system may be, include, or otherwise be included in a mobile device (e.g., smartphone, tablet computer, laptop computer, computerized watch, computerized eyewear,
computerized gloves), a personal computer, a smart television, a personal digital assistant, a portable gaming system, a media player, a mobile television platform, an automobile navigation and/or entertainment system, a vehicle (e.g., automobile, aircraft) and/or cockpit display, a home or other smart appliance and/or related device (e.g., interconnectable appliance/device via Internet of Things), or any other type of wearable, non-wearable, mobile, or non-mobile computing device, and the computing system may or may not include a display device. In some cases, the interactive assistant may be a voice-assistant that receives audible user commands, processes the commands based on speech recognition operations, and performs corresponding actions, such as providing audible responses to user queries and/or performing certain actions. The interactive assistant may provide or utilize a user interface with which a user can communicate to cause the assistant to output useful information, respond to a user’s queries, or otherwise perform certain operations to help the user complete a variety of real-world or virtual tasks.

Figure 1 below illustrates an example of an interactive assistant that may to dynamically adjust the playback of content (e.g., audible content output by the interactive assistant) in response to receiving a user command. As noted above, a user may engage in various different interactions with the assistant by submitting queries, performing searches, requesting select information, providing instructions, and the like. At any point in time, during playback of content provided by the interactive assistant, the user may issue a command to adjust the playback of such content. For example, the user may issue an audible command to instruct the interactive assistant to pause the playback of content, and may later issue a subsequent command to resume or continue playback of the content. The user may also issue a command that causes the interactive assistant to skip over certain playback content or to repeat certain content that was
previously output. In such fashion, the interactive assistant is thereby capable of dynamically navigating to various points in a content playback timeline.

Figure 1

As shown in Figure 1, the computing system that includes the interactive assistant may have or otherwise be communicatively coupled to one or more input devices and one or more
output devices. For instance, the input devices may include one or more microphones, a presence-sensitive input device and/or touch-sensitive screen, a mouse, a keyboard, a voice responsive system, a camera, or any other type of device for detecting input from a human or machine. In some cases, the input device may one or more location sensors (GPS components, Wi-Fi components, cellular components), one or more temperature sensors, one or more movement sensors (e.g., accelerometers, gyroscopes), one or more pressure sensors (e.g., barometer), one or more ambient light sensors, and/or one or more other sensors (e.g., camera, infrared proximity sensor, hygrometer, and the like). Other sensors may include a heart rate sensor, magnetometer, glucose sensor, hygrometer sensor, olfactory sensor, compass sensor, step counter sensor, to name a few other non-limiting examples.

The computing system may also include or be communicatively coupled to one or more output devices, such as one or more speakers or display screens, including a presence-sensitive screen and/or a touchscreen, or any other type of device for generating output to a human or machine. In some cases, the input devices and/or output devices may include one or more other type of wearable, non-wearable, mobile, or non-mobile computing devices that are also used by the user. One or more of the input and/or output devices may be external to and communicatively coupled (e.g., via a wired or wireless connection) with the computing system.

The computing system may also include a user interface module that is configured to manage inputs received by the interactive assistant as users interact with the computing system, and the user interface module may be configured to receive additional instructions from applications, services, platforms, or other modules of the interactive assistant that process user input. The user interface module may also be configured to process output that is provided to users, and may be coupled to the input device(s) and output device(s) of the interactive assistant.
The computing system may also include a speech recognition module, which may interface with the user interface module and/or the interactive assistant. When a user provides audible input to the interactive assistant (e.g., via commands, questions, queries), the interactive assistant may use the speech recognition module to process such audible input.

As described herein, the speech recognition module, user interface module, and interactive assistant are configured to utilize and/or process information received from the input devices only after receiving explicit authorization from the user to do so. The computing system may provide the user (e.g., via the user interface module and/or the interactive assistant) with detailed information about the requested use of data collected by input devices for use by the interactive assistant, in conjunction with the use of specified functions or applications (e.g., one or more of the search application, content delivery applications, or other applications). The speech recognition module, user interface module, and/or interactive assistant only use such data collected by input devices after receiving explicit authorization from user to do so. After receiving such authorization, these modules are configured only to use the data gathered by the input devices for the use or uses authorized by the user.

Further to the descriptions below, a user may be provided with controls allowing the user to make an election as to if and when the interactive assistant, the computing system, and/or the external computing systems described herein can collect or make use of supplemental data (e.g., user information or contextual information about a user’s social network, social actions or activities, profession, a user’s preferences, or a user’s current location), and if and when the user is sent content or communications from a computing system. In addition, certain data may be treated in one or more ways before it is stored or used, so that personally identifiable information is removed. For example, a user’s identity may be treated so that no personally identifiable
information can be determined for the user, or a user’s geographic location may be generalized where location information is obtained (such as to a city, ZIP code, or state level), so that a particular location of a user cannot be determined. Thus, the user may have control over what supplemental data is collected about the user, how that supplemental data is used, and what supplemental data is provided to the user.

As shown in Figure 1, the computing system includes a search application. Over time, and after having provided the explicit authorization noted above to approve the processing of user input data, the user may submit one or more queries to the interactive assistant to request content. Based on such queries, the search application may search for content in one or more content repositories that are local to the computing system and/or stored on one or more external computing systems (e.g., content providers), such as shown in Figure 1. Upon identifying content that is responsive to the user’s queries (e.g., content that satisfies one or more criteria specified in the queries), the search application may output the identified content to the user, such as via audible and/or visual output. In addition, the search application may store the history of these searches in one or more datastores containing history information. As shown in Figure 1, the history information may be stored locally on the computing system and/or on the one or more external computing systems, and may include any information or metadata associated with these searches and corresponding search results. This history information may be part of or otherwise included in the contextual information that is used by one or more content delivery applications.

The search application and the content delivery applications may be executed locally on the computing system. In some cases, however, any of the search application or the content
delivery applications may be executed on the one or more external computing systems and may be invoked by the interactive assistant.

The content delivery applications may include one or more applications that are configured to use such contextual information and deliver different types of content to the user. For instance, the content delivery applications may include one or more applications to deliver news content, video content, image content, map content, audiobook content, and the like. When the content is included in the content repositories local to the computing system, the content delivery applications and/or the interactive assistant may output the content directly via the output devices. When the content is included in content repositories stored on the one or more external computing systems, the external computing systems may stream the content to the computing system, and the computing system may then output the received content via the output devices. In such fashion, the computing system is capable of delivering received content even if the computing system is, at a certain point in time, in an offline mode or otherwise disconnected from the external computing systems.

At any point in time, during playback of content provided by the interactive assistant, the user may issue a command to adjust the playback of such content. For example, the user may issue an audible command to instruct the interactive assistant to pause the playback of content, and may later issue a subsequent command to resume or continue playback of the content from the point at which playback was previously paused. By issuing a pause or similar command, the user may cause the interactive assistant to at least temporarily stop or pause the playback of content. For instance, as one non-limiting example, the user may have previously requested instructions from the interactive assistant to cook and prepare a particular meal or food item. In response, the interactive assistant and/or search module may identify cooking instructions
associated with a recipe in one or more content repositories located on the computing system
and/or on the external computing systems. The content delivery applications may then prepare
such content for output to the user via the interactive assistant and/or user interface module.

However, during playback (e.g., audible output) of these cooking instructions to the user,
the user may wish to temporarily pause playback in order to complete a particular cooking
instruction, to obtain additional food items from the refrigerator, or other reasons. In this case, as
shown in the example of Figure 1, the user may issue a command to “pause playback” (or,
alternatively, e.g., to “stop playback,” “hold on,” “hang on,” “wait,” or “I’m not there yet,” to
name a few other non-limiting examples). In this case, the interactive assistant and/or content
delivery applications may temporarily stop or pause the playback of these cooking instructions.
The interactive assistant and/or content delivery applications may identify the point in time of the
content at which playback was paused, which may comprise a particular point in the content
playback timeline associated with the cooking instructions.

For instance, the entire set of cooking instructions may be 5 minutes in length. If the user
issues the command to “pause playback” half-way through the instructions, the interactive
assistant and/or content delivery applications may identify the half-way point of the content, or 2
minutes and 30 seconds from the start of content playback in the overall playback timeline.

Once the user is ready to resume playback, the user may issue a command to continue or
resume playback (e.g., a user command to “continue playback,” “resume playback,” “okay, keep
going,” “okay, I’m ready,” “next,” or “what’s next,” to name a few non-limiting examples). At
this point, the interactive assistant and/or content delivery applications may resume playback of
the content at the point at which playback had been previously stopped or paused (e.g., from the
half-way point of the content, or 2 minutes and 30 seconds from the start of content playback).
Continuing with this example, at a certain point during playback of the cooking instructions, the user may issue a command that causes the interactive assistant to skip over certain playback content or to repeat certain content that was previously output. In such fashion, and based on user feedback, the interactive assistant and/or content delivery applications may be capable of dynamically navigating to various points in a timeline of content associated with the cooking instructions played back or otherwise output to the user.

For instance, as shown in Figure 1, the user may issue a command to “skip forward to next instruction” (or, e.g., “skip to next one,” “move to next instruction,” or “go to next instruction,” to name a few other non-limiting examples). In this case, perhaps the user is familiar with the portions of the cooking instructions that are currently being played back (e.g., based on prior use of the instructions by the user), but the user cannot recall the next set of instructions and wishes to skip forward to the next instruction of the instruction content.

Upon receiving this audible command, the interactive assistant and/or content delivery applications may skip forward to the next instruction in the playback timeline of the instruction content. In this case, the content delivery applications may access information associated with the content, such as metadata, which identifies specific locations in the playback timeline that are associated with corresponding cooking instructions. This information may be stored in one of the content repositories locally on the computing system or on the external computing systems. Using this information, the content delivery application is capable of determining the particular starting location in the playback timeline that is associated with the start of the next instruction requested by the user, traversing to this location in the timeline, and playing back the next instruction for output to the user.
Similarly, the user may issue a separate command to “repeat last instruction” (or, e.g., “what was that,” “I missed that,” “say instruction again,” to name a few other non-limiting examples), such as if the user didn’t fully understand or hear the current instruction and wants to hear again. In this case, the content delivery applications may determine the particular starting location in the playback timeline that is associated start of the current or last played instruction, traversing to this location in the timeline, and repeating the instruction for the user. The content delivery applications may continue with output of the content from that point forward unless the interactive assistant receives another command from the user to traverse to another portion of the content.

In such fashion, the interactive assistant and/or the content delivery applications may be configured to process skip and repeat commands from the user to respectively add or subtract units of time with respect to the playback timeline in determining which portions of content to output to the user. Furthermore, in similar fashion, in some alternate cases, the interactive assistant may process direct content traversal commands, such as “fast forward” and “rewind” commands from the user, which specify explicit units of time. For example, the user may issue a command to “fast forward 2 minutes” (or, e.g., “go forward 2 minutes,” “go ahead 2 minutes,” “forward 2 minutes,” to name a few other non-limiting examples) to cause the interactive assistant and/or content delivery applications to skip forward in the content playback content 2 minutes, or may issue a command to “rewind 2 minutes” (or, e.g., “go backward 2 minutes,” “go back 2 minutes,” “back 2 minutes,” to name a few other non-limiting examples) in order to go backwards in the content playback content 2 minutes (e.g., if a user is watching some form of video content provided by the content delivery applications).
In some instances, the user may issue a command to the interactive assistant that causes the interactive assistant to output the cooking directions one instruction at a time. The user’s command may explicitly request that the assistant output the instructions one at a time, or the assistant may automatically perform such functionality even without an explicit user request (e.g., when the content includes itemized instructions or other information). The interactive assistant may output the first instruction for the cooking recipe and then pause to wait for the user’s feedback to proceed to the next instruction. The interactive assistant may repeat this procedure for each cooking instruction in the set of cooking instructions, and the assistant may even generate and output interstitial language to remind the user about the pending state (e.g., a particular location or instruction number within the set of cooking instructions, such as, for instance, the “third step”).

For instance, the user may issue a command to “read me the directions for the chocolate chip cookie recipe.” In response, the interactive assistant may output the first cooking instruction for the recipe (e.g., “okay, start by pre-heating the oven to 350 degrees”). After outputting this first instruction, the interactive assistant may further output an indication that it is awaiting further user feedback before proceeding to the next instruction (e.g., “I will wait until you say ‘next step’”). Once the user has pre-heated (or at least begun pre-heating) the oven to 350 degrees and is ready for the next instruction, the user may say provide such indication to the assistant (e.g., the user may say “next step”). In response, the assistant may output the next cooking instruction, as well as an indication that it is awaiting further user feedback or confirmation before proceeding to the next instruction (e.g., “okay, take out one egg and three cups of flour… I’ll wait until you’ve done so. Just say next step when you’re ready”).
Once the user is ready, the user may provide such indication or confirmation to the assistant (e.g., “next step please”). In some cases, however, the user may wish for the assistant to repeat the last instruction (e.g., “please repeat”) or may even issue a question. For instance, the user may ask, “How many eggs was that?” In response, the assistant may output the answer of “one egg.” In this fashion, the user is able to control content playback by the assistant, not necessarily by explicitly stopping or interrupting playback but instead by progressing it.

As another non-limiting example, the user may request that the interactive assistant provide audio output of a specified audiobook. In response to receiving the user’s request, the content delivery applications may prepare content of the audiobook for output to the user via the interactive assistant and/or user interface module.

At any point during playback of the audiobook, the user may wish to temporarily pause playback. In this case, similar to the example shown in Figure 1, the user may issue a command to “pause playback” (or, alternatively, to “stop playback” or to perform one of the other example commands noted above). In this case, the interactive assistant and/or content delivery applications may temporarily stop or pause the playback of the audiobook and identify the point in time of the content at which playback was paused, which may comprise a particular point in the content playback timeline associated with the audiobook. Once the user is ready to resume playback, the user may issue a command to continue or resume playback (e.g., a user command to “continue playback” or “resume playback” or to perform one of the other example commands noted above). At this point, the interactive assistant and/or content delivery applications may resume playback of the audiobook content at the point at which playback had been previously stopped or paused.
Based on user feedback, the interactive assistant and/or content delivery applications may be capable of dynamically navigating to various points in a timeline of content associated with the audiobook. For instance, similar to the example shown in Figure 1, the user may issue a command to “skip to next chapter” (or, e.g., “move to next chapter,” “forward to next chapter,” or “go to next chapter,” to name a few other non-limiting examples) Upon receiving this audible command, the interactive assistant and/or content delivery applications may skip forward to the next chapter in the playback timeline of the instruction content. The content delivery applications may access information associated with the audiobook content, such as metadata, which identifies specific locations in the playback timeline that are associated with corresponding chapters. Using this information, the content delivery application is capable of determining the particular starting location in the playback timeline that is associated with the start of the next chapter requested by the user, traversing to this location in the timeline, and playing back the next chapter for output to the user. Similarly, the user may issue other navigation commands to navigate to other portions in the timeline of the audiobook content (e.g., “skip forward to next page,” “skip forward to next section,” “repeat last page,” “repeat last chapter”).