November 2020

Picking Best Suited Smart Device to Provide Response to a Query

Michael Schäer

Agoston Weisz

Vitaly Gatsko

Picking Best Suited Smart Device to Provide Response to a Query

ABSTRACT

Users can issue commands, e.g., spoken commands, to virtual assistants available on various smart devices that have different capabilities. Responses to different types of queries may be best served on specific device types. This disclosure describes techniques to provide responses to a user query, e.g., a spoken query to a virtual assistant, via an automatically selected device from available user devices. Such selection can be based on explicit user preferences or can be based on user-permitted data, as well as be based on the type of query. Implementation of the described techniques can provide an enhanced user experience of interaction with a virtual assistant.

KEYWORDS

- Device arbitration
- Virtual assistant
- Spoken query
- Smart speaker
- Smart display
- Smart appliance

BACKGROUND

Users can issue commands, e.g., spoken commands, to virtual assistants available on various smart devices such as smartphones, smart speakers, smart displays, or other appliances. The device can provide a response to the command via various types of feedback channels, e.g., an audio response, a displayed response, etc. Different devices that include virtual assistant functionality have different capabilities. For example, smart speakers and headphones may have good audio playback quality but lack displays and input capability other than voice; smartphones

include touchscreen displays, but provide relatively low quality audio playback and are inferior to larger screens when playing video; etc.

Many users have multiple devices that all implement the same virtual assistant application. Any such device can receive a command, e.g., a spoken command and provide a response. For example, a user may have a smart speaker and a smartphone, both in the same room at a time of a spoken command; a user may have a smartphone and connected wireless headphones; etc. In these situations, where multiple devices may include different capabilities, and may receive a user command, a decision needs to be made on which smart device(s) the response to the query is served. When multiple devices recognize the input query, depending on the query, the device that is best-suited for providing the response may vary. For example, audio responses may be well-served by a smart speaker, while displayed responses may be better served on a device with a screen, e.g., a smartphone or smart display.

DESCRIPTION

This disclosure describes techniques to automatically determine the best-suited virtual assistant device to provide a response to a user query. Fig. 1 illustrates an example scenario of a room (100) that has multiple smart devices. A smart speaker (108) and a smartphone (110) both include virtual assistant capability, e.g., enable users to provide spoken commands to which the device responds.

Fig. 1(a) shows an example query (104) where the user (106) asks the virtual assistant via an audio command, e.g., that includes a wake word to invoke the virtual assistant, followed by a command "play 80's rock." Since the appropriate response to the query is to initiate playback of 80's rock music, the response (102) is determined as best served on the smart speaker (108) (which has better audio playback capability than the smartphone).

Fig. 1 (b) illustrates another user query (112) in the same room - the query is about soccer standings for a specific tournament. While results can be delivered via audio, it is likely that soccer standings are better presented as a table (e.g., which can include a large number of teams, show other information such as total points scored, goals for/against, recent results, etc. Therefore, it is determined the virtual assistant response (114) is likely best served on a device with an interactive display, which in this case is the smartphone. The user can then further interact with the results and view additional information as needed. Thus, even though it was possible to provide a response to the query as audio via the smart speaker, a better quality user experience is provided by providing the response via the smartphone.

While Fig. 1 shows examples of responses on individual devices, multiple devices can be used together to provide the response. For example, while audio playback of 80's rock is provided via the smart speaker, lyrics may be displayed automatically on the smartphone. Also, the smart speaker can provide a recitation of the top 3 soccer teams and indicate to the user that a table with the current standings is available via their smartphone. While these example relate to audio and displayed responses, other types of media, e.g., videos (best-suited on a large screen device), vertical videos (better suited on a vertical display such as a smartphone), navigation guidance (best-suited on a mobile device which can be carried by the user, rather than a fixed stationary device), augmented reality/ virtual reality experiences (best provided via a device with such capabilities), can also be included in the response and appropriate devices can be picked automatically. The user can also be provided an option to receive a follow-up response on another device. For example, if the user has a phone paired with wireless headphones and asks for popular websites about a particular topic, the URL can be mentioned via the headphones and a follow-up option may be provided, e.g., "would you like to open this link on the phone?"
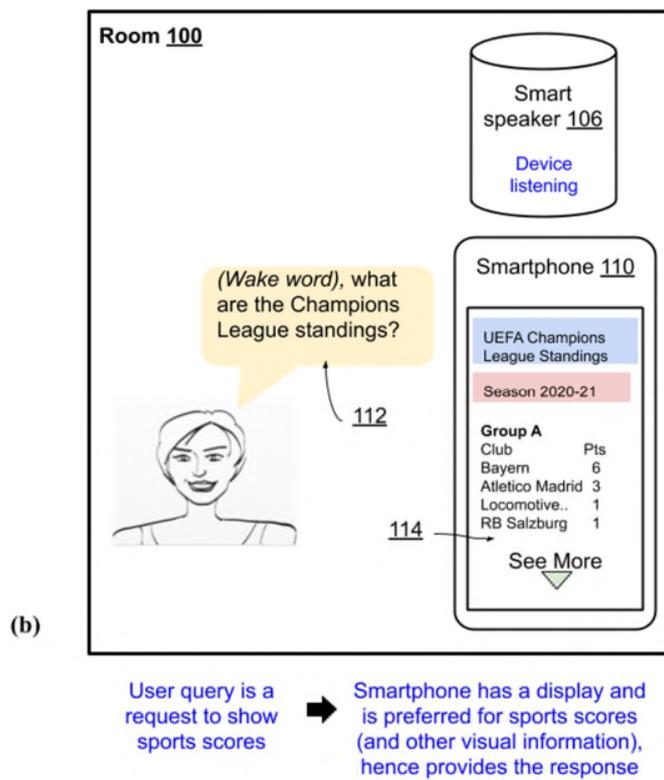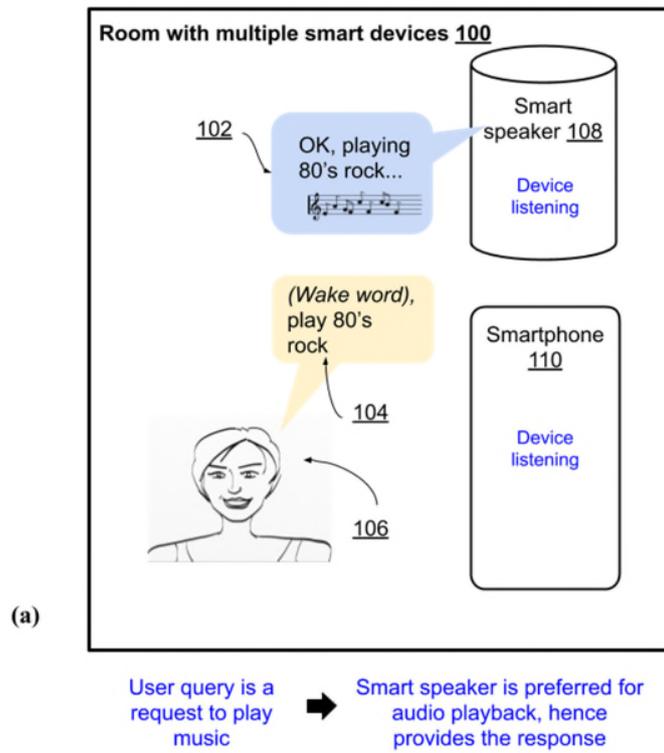
Fig. 1: (a) speaker device preferred for audio; (b) smartphone preferred for GUI display

The techniques of this disclosure can be used in combination with techniques that detect the device that the user is addressing when issuing a query. The techniques can automatically detect the types of queries that are best served via a particular device or type of device. An important parameter in such determination is the respective capability of each available device. For example, when a smart speaker (without a display) and a smartphone is available, and the query relates to images ("show me my photos"), the smart speaker is determined to be inappropriate and the response is served via the smartphone. In another example, with user permission, it may be detected that a query for music playback was received when the user is home where a smart speaker is available, or in their car, that has an in-car audio system. In this case, the smart speaker or in-car audio system is selected for music playback owing to the superior audio quality.

The selection of a particular device for providing response to a user query may be based on user-specific factors, if use of such factors is permitted by the user. For example, with user permission, prior queries and responses, and interactions with the virtual assistant can be analyzed to select the device that provides the response. Such analysis can include, e.g., determining whether the current query is: same as a prior query; semantically similar to prior queries; or related to a specific feature, e.g., setting a timer. Suitable heuristics, rule-based engines, or machine learning techniques can be utilized, with user permission, to perform such analysis. With permission from the users, aggregate information about user preferences can be analyzed and used to select the device to provide a response, e.g., users generally prefer televisions for playback of high definition content when a tablet device and smart television are both available.

The analysis may be based on explicit user input and/or implicit user feedback, inferred from user actions. Factors that serve as input (with user permission) for analysis algorithms can include one or more of:

- **User interaction with a virtual assistant response:** Often, devices with a feedback channel (e.g., a display screen) include input capabilities allowing users to interact on that channel (e.g., a touchscreen). The more interaction a user provides on the feedback from a channel, the more likely it is that devices that provide such a channel are preferred. In particular, if users show substantial levels of interaction with a response provided on devices with a touchscreen, e.g., scrolling to view more images, clicking on links, etc., it is a strong indicator of user preference for devices with touchscreen capabilities.

- **Ratio of repeating the query on a second device:** If the user issues the same (or similar) query on a second device immediately after receiving a virtual assistant response, it can be an indication that the user prefers the second device rather than the device that initially provided the response. For example, if the user asks for how an animal sounds, the user may be satisfied with a response from a smart speaker (no follow-up queries). while if the user requests videos of the animal making sounds, follow-up queries are received via a smart display (or other device that has a screen), indicating that the second type of query is better served by devices with screen capabilities.

- **Query stream distribution across different device types:** Comparing the query volume of different device types with channels of varied feedback/capabilities can be indicative of user preference. A query type has a much higher ratio of occurrence on a particular device type than the overall average may be determined as likely better served on such device types. For

example, navigation queries are much more frequent on mobile devices than devices mounted at home, and may thus be determined as best served via mobile devices.

Further to the descriptions above, a user is provided with controls allowing the user to make an election as to both if and when systems, programs, or features described herein may enable collection of user information (e.g., information from sensors of a user device, smart devices that a user accesses, a user's prior queries and interactions with a virtual assistant, a user's current location), and if the user is sent content or communications from a server. In addition, certain data are 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 is treated so that no personally identifiable information can be determined for the user. Thus, the user has control over what information is collected about the user, how that information is used, and what information is provided to the user.

CONCLUSION

Users can issue commands, e.g., spoken commands, to virtual assistants available on various smart devices that have different capabilities. Responses to different types of queries may be best served on specific device types. This disclosure describes techniques to provide responses to a user query, e.g., a spoken query to a virtual assistant, via an automatically selected device from available user devices. Such selection can be based on explicit user preferences or can be based on user-permitted data, as well as be based on the type of query. Implementation of the described techniques can provide an enhanced user experience of interaction with a virtual assistant.

REFERENCES

1. Leong, Jian Wei. "Virtual assistant identification of nearby computing devices." U.S. Patent 10,332,523, issued June 25, 2019.