March 2020

Automatic Input Mode Selection for Text Entry

Vilem Ries

Automatic Input Mode Selection for Text Entry

ABSTRACT

Users can provide input to a computing device via various modalities such as keyboard, mouse, via voice, etc. When interacting with a virtual assistant, users currently need to switch manually between keyboard and voice input. Such cumbersome interaction can lead to reduced usage of a virtual assistant or other software. This disclosure describes a user interface in which a virtual keyboard and voice input mode are both activated simultaneously. When the user starts providing input via one modality, the other modality is automatically turned off. The user interface enables fast interaction with a virtual assistant or other software.

KEYWORDS

- Input mode
- Voice input
- Keyboard input
- Multimodal input
- Virtual assistant

BACKGROUND

Speed and efficiency of interaction are important characteristics of a user interface. A user interface that is cumbersome can cause a user to lower their usage of the corresponding product, such as a virtual assistant. For example, to interact with a virtual assistant, in some situations, a user prefers to provide input via a virtual keyboard, while in other scenarios, the user prefers to provide spoken input. Current virtual assistant UI requires that the user select the desired input mode prior to providing the input, thus creating fiction.

DESCRIPTION

This disclosure describes a user interface in which a virtual keyboard and voice input mode are both activated simultaneously. When the user starts providing input via one modality, the selected modality is enabled and the other modality is automatically turned off and removed from the displayed user interface. The user interface enables fast interaction with a virtual assistant or other software.

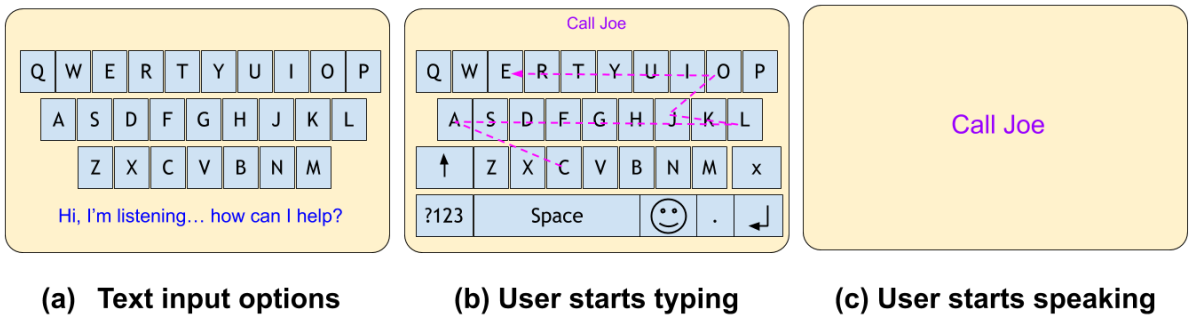

Fig. 1: Multimodal user interface for text entry

Fig. 1 illustrates a multimodal user interface for text entry. As illustrated in Fig. 1(a), the user interface includes an element - "Hi, I'm listening… how can I help?" - that indicates that a device microphone is active and a virtual assistant application is ready to accept spoken input. The interface also includes a simplified virtual keyboard, e.g., that includes text characters but excludes other keys such as space, backspace, emoji selection, shift, etc. The device microphone is activated automatically with specific user permission.

As illustrated in Fig. 1(b), if the user starts typing, as indicated by the swipe pattern, the element indicating availability of voice input is removed from the interface. The virtual keyboard is updated to a full-featured keyboard that includes a more complete set of keys. The user can continue typing to provide text input, e.g., "Call Joe."

As illustrated in Fig. 1(c), if the user starts speaking, the virtual keyboard is removed from the keyboard and is replaced by a text field that displays a transcription of the spoken input, e.g., “Call Joe.”

By automatically switching modes between keyboard or voice input based on the initial action performed by the user, the user interface as described herein eliminates the need for the user to make an explicit mode selection prior to providing input, thus reducing friction. Spoken input is analyzed using speech recognition techniques. Once the input is complete, the virtual assistant can perform an action corresponding to the command, e.g., place a call to another user named Joe.

While this disclosure describes a virtual assistant with a user interface that simultaneously displays input modes of voice and keyboard, other input modes can also be displayed and combined on the user interface, such as touch, gesture, pen, etc. The techniques can be implemented to provide input to any software application. The user is provided with options to turn off certain modalities (e.g., voice) and in such configuration, the user interface is updated to show only the modalities that are available.

Further to the descriptions above, a user may be 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 about a user’s social network, social actions or activities, profession, a user’s preferences, or a user’s current location), and if the user is sent content or communications from a server. 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 information is collected about the user, how that information is used, and what information is provided to the user.

CONCLUSION

This disclosure describes a user interface in which a virtual keyboard and voice input mode are both activated simultaneously. When the user starts providing input via one modality, the other modality is automatically turned off. The user interface enables fast interaction with a virtual assistant or other software.