January 2022

Facilitating Discovery of Virtual Assistant Capabilities via Augmented Reality

Eric Stavarache

Felix Weissenberger

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

Facilitating Discovery of Virtual Assistant Capabilities via Augmented Reality

ABSTRACT

Users may not be aware of all possible tasks that a virtual assistant can perform. This disclosure describes techniques for user-permitted use of augmented reality (AR) technology to help users discover contextually relevant tasks that a virtual assistant could perform. With user permission, the relevant tasks are determined by recognizing entities and/or locations within the camera view of the user's device. Additional data from device sensors and/or other known user information such as devices owned or apps installed can also be utilized with user permission. The contextually relevant tasks are recommended to the user as an augmented message, via audio, or other interfaces, thus helping the user discover virtual assistant capabilities.

KEYWORDS

- Virtual assistant
- Voice-based assistant
- Augmented Reality (AR)
- Feature discovery
- Feature discoverability

- Task recommendation
- Smart appliance
- Smart home
- Object recognition
- Computer vision

BACKGROUND

Users often rely on a virtual assistant to handle a variety of tasks, such as placing food orders, scheduling cab rides, arranging taxi trips, setting alarms, jotting notes, controlling devices within a smart home, etc. Having a virtual assistant perform such tasks provides a convenient user experience (UX) and can boost productivity.

A virtual assistant can be used for performing tasks supported by augmented reality (AR). AR typically involves providing relevant augmented capabilities in relation to objects and/or

locations recognized to be present in images or videos of the real world captured by an AR device. For instance, a user can point the device camera at a restaurant menu to invoke AR capabilities such as viewing the menu translated to another language. However, tasks supported by a virtual assistant are not easily discoverable. Therefore, users may not be aware of all possible tasks that a virtual assistant can perform on their behalf. Such lack of discoverability can result in users being burdened to perform tasks that could have easily been delegated to a virtual assistant, which is inconvenient and inefficient.

DESCRIPTION

This disclosure describes techniques for the use of AR technology in combination with a virtual assistant to recommend tasks that the virtual assistant could perform for a user. The recommendations, generated and provided with user permission, can help users discover various tasks relevant to the context that a virtual assistant could perform on their behalf.

The recommendations for potentially relevant tasks supported by the virtual assistant are based on pictures and/or videos of the user's view from an augmented reality device, obtained from the device with permission of the user. The pictures and/or videos of the user's view are analyzed using a suitably trained machine learning model to recognize objects (e.g., vacuum cleaner, lamp, map, street sign, etc.) and/or locations (e.g., home, street, gym, forest, etc.). For instance, when a user is looking at a vacuum cleaning robot at home, a virtual assistant can suggest, "Try saying '*Vacuum the living room*' to activate the cleaning robot." Similarly, the virtual assistant can recommend tasks for turning on the lights when looking at a lamp, navigating home when the user is looking at a map, etc.

Data from device sensors, obtained with user permission, can be combined with the information from the user's view when determining the relevant task to be suggested. For

example, if the user is on a hiking trail within a forest and moving up, the virtual assistant can

suggest that the user can issue the command "*enter hiking mode*". Alternatively, or in addition,

the recommendation can be derived by combining information in the user's view with other

known user information accessed with user permission. Such information can include, for

example, devices owned, apps installed, etc. For example, if the user is looking at a vacuum

cleaner, the suggested command can be customized to the specific brand of the vacuum cleaner

the user is known to own and the specific version of the home control app the user has installed

on the device.
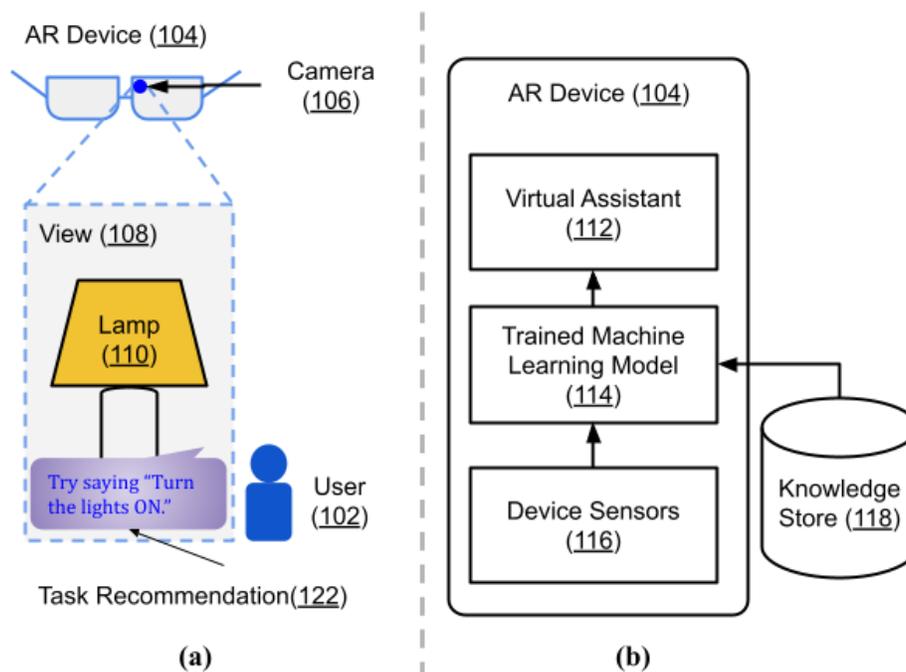
Fig. 1: Recommending relevant tasks for a virtual assistant based on AR view

Fig. 1 shows an operational implementation of the techniques described in this disclosure.

As seen in Fig. 1(a), Aa user (102) captures an image of a lamp (110) within the user's view

(108) using the camera (106) on the user's device (104), e.g., augmented reality glasses,

smartphone, or any other device. As seen in Fig. 1(b) which depicts operation to generate a

suggested command, the captured picture of the view is input to a trained machine learning

model (114) along with data from the device sensors (116) and a knowledge store (118) of

known information, such as the user's devices and apps, commands supported by a virtual

assistant, etc. The output of the model includes one or more recommended commands that the

user could issue to a virtual assistant to perform tasks relevant to the user's current view. As seen

in Fig. 1(b), the recommendation (122) is delivered to the user by a virtual assistant (112) as an

augmentation within the user's view. Alternatively, or in addition, the recommendation can also

be provided via a speaker of the device.

The techniques described above can be invoked automatically whenever the user is in AR

mode. Alternatively, the user can manually request the recommendations from the virtual

assistant using any suitable interaction mode, such as a voice command, graphical user interface

(GUI), etc. Similarly, task recommendations can be delivered to the user via any suitable

mechanism, such as speech, text, etc.

With user permission, the techniques can be implemented within any device, application,

or platform that supports AR capabilities. If users permit, the techniques can be applied to

discover tasks that can be delegated to any virtual assistant application or service for entities

within the user's physical environment that support the use of virtual assistants. The

recommendations for relevant tasks that can be performed by a virtual assistant can be obtained

from a knowledge store of available commands maintained by developers of virtual assistants

and/or provided by developers of apps connected to specific devices, locations, or activities.

Implementation of the techniques can help users utilize the AR capabilities of their

devices to discover device and app features relevant to their physical environments and activities

that can be delegated to a virtual assistant. The techniques leverage AR seamlessly to enhance

the user experience and effectiveness of using a virtual assistant, thus boosting the user's efficiency and productivity.

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 devices, objects in the user's view, 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 techniques for user-permitted use of augmented reality (AR) technology to help users discover contextually relevant tasks that a virtual assistant could perform. With user permission, the relevant tasks are determined by recognizing entities and/or locations within the camera view of the user's device. Additional data from device sensors and/or other known user information such as devices owned or apps installed can also be utilized with user permission. The contextually relevant tasks are recommended to the user as an augmented message, via audio, or other interfaces, thus helping the user discover virtual assistant capabilities.

REFERENCES

1. Badr, Ibrahim et al. "Identifying and Controlling Smart Devices." U.S. Patent Application

    Publication No. 20200319765A1, filed June 18, 2020.

2. Badr, Ibrahim. "Selecting an input mode for a virtual assistant." U.S. Patent Application

    Publication No. US20190354252A1, filed May 16, 2018.