May 11, 2018

Interaction with smart assistants using alternative and augmentative communication

Lorenzo Caggioni

Paolo Pigozzo

Davide Ferraro

Egil Hogholt

Teresa Pato

See next page for additional authors

Inventor(s)
Lorenzo Caggioni, Paolo Pigozzo, Davide Ferraro, Egil Hogholt, Teresa Pato, and Thomas Riga

Interaction with smart assistants using alternative and augmentative communication

ABSTRACT

Smart assistants utilize speech recognition, sensing, artificial intelligence, and networking technologies to enable improved human-machine interaction. However, there remain use cases where smart assistants are not easily usable by humans. For example, voice-activated assistants are not accessible to those that are hearing or speech impaired. Touchscreen based assistants are not usable by those who lack fine motor skills and/or reading ability.

This disclosure adds to the modalities by which humans can control and communicate with smart assistants by enabling use of physical objects, facial expressions, gross motor skills, body movements, etc. to provide commands. Collectively, these techniques of control and communication are referred to as alternative and augmentative communication (AAC).

KEYWORDS

- Alternative and augmentative communication
- AAC
- smart assistant
- accessibility
- gesture recognition

BACKGROUND

Smart assistants utilize speech recognition, sensing, artificial intelligence, and networking technologies to enable improved human-machine interaction. However, there remain use cases where smart assistants are not easily usable by humans. For example, voice-activated assistants are not accessible to those that are hearing or speech impaired. Touchscreen based assistants are not usable by those who lack fine motor skills and/or reading ability.

Accessing the Internet and using devices such as a television, smartphone, thermostat, home automation devices, etc. can be difficult for differently-abled individuals, since interaction with such devices requires some fine-motor skill or fluency with speech, hearing, reading, etc. Senior citizens, young children that are yet to develop fine motor skills, and others also face similar difficulties using computer user interfaces.

DESCRIPTION

This disclosure expands the numbers and types of modalities or channels of human-machine communication by associating actions such as body movements, gestures, facial expressions, gross motor movements, physical objects, etc. with commands. A simple method is provided for a conventionally-abled caregiver to associate an action with a command. The action-command association is reconfigurable. At a given time, a unique action is associated with a unique command. The techniques of this disclosure are illustrated with examples.

The techniques described herein utilize user input such as movement of objects, facial expressions, button presses, etc. with user permission. Users are provided with options to change or disable recognition of such input, e.g., by turning off one or more sensors of a device that implements a smart assistant, e.g., a camera, a microphone, a proximity sensor, etc. Recognition of user input is performed using only the sensors for which the user has provided permission.

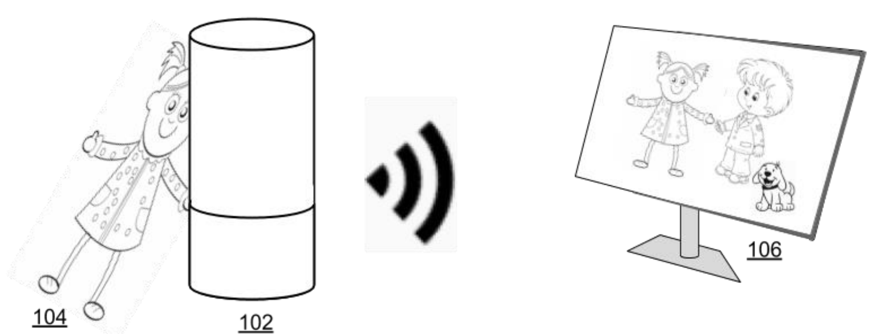

Fig. 1: Smart assistant activated by the presence of a particular physical object

Example 1: Physical object that serves as a command. In Fig. 1, a smart assistant available via a device (102) is configured such that the presence of a particular doll (104) causes television (106) to launch an animation channel that features the doll. The doll is detected by the smart assistant, e.g., using techniques such as computer vision, proximity sensing, near-field communications, RFID tagging, Bluetooth beacons, etc. A different physical object, e.g., a doll of a different animation character, can be configured to launch a different movie or TV channel.

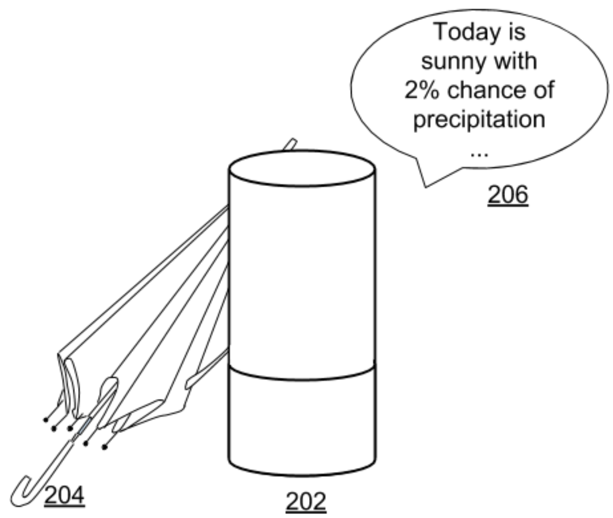

Fig. 2: Smart assistant activated by the presence of a particular physical object

Example 2: Physical object that serves as a command. In Fig. 2, a smart assistant available via a device (202) is configured such that the presence of an umbrella (204) proximate to the device causes the device to recite the weather prediction (206) for the day.

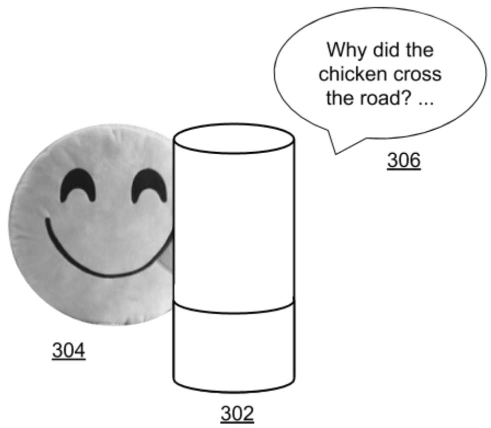

Fig. 3: Smart assistant activated by the presence of a particular physical object

Example 3: Physical object that serves as a command. In Fig. 3, a smart assistant available via a device (302) is configured such that the presence of a smiley-face pillow (304) in proximity to the device causes the device to recite a joke.

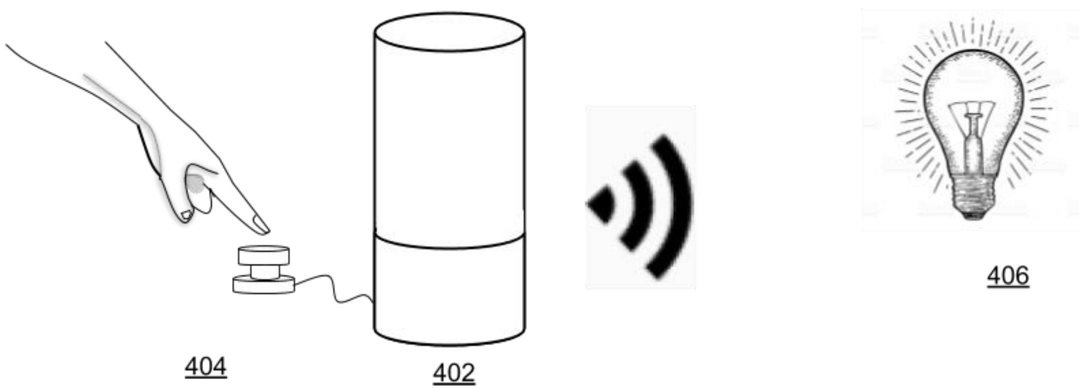

Fig. 4: Smart assistant activated by gross motor movement

Example 4: Gross motor movement serving as command. In Fig. 4, a smart assistant available via a device (402) is configured such that a gross motor movement, e.g., the pressing of a button (404), causes activation of another device, e.g., bulb (406) switching on. For example, the device that is activated can be an internet-of-things (IoT) device, e.g., thermostat, television, etc.

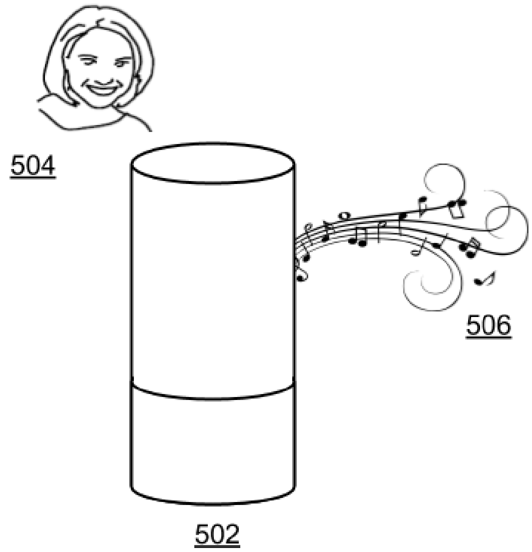

Fig. 5: Smart assistant activated by facial expression

Example 5: Facial expression or gesture serving as command. In Fig. 5, a smart assistant available via a device (502) is configured such that detection of a smiling human face (504) causes the device to start music playback (506). Recognition of the expression is performed only when users permit such recognition. Users are provided with options to disable the recognition. In a similar manner, the detection of a frowning human face causes the device to stop music playback. The musical piece can be pre-configured, e.g., by a caregiver, or tailored based on user preferences of the user that provided the expression or gesture as input.

The input techniques of this disclosure can be used by all types of users, including differently-abled individuals, senior citizens, or young children. The techniques are implemented upon permission from a consenting user, e.g., a parent, a caregiver, etc.

Example 6: Body movement serving as command. A conventionally-abled individual has a certain bedtime routine, e.g., locking all doors, turning down the thermostat, setting the alarm, etc. With user permission, a smart assistant device is configured to automatically perform the bedtime routines upon detecting that the individual has retired to bed.

<u>Example 7: Physical object serving as command.</u> A smart assistant device is configured such that when a user (e.g., a parent) points to a nearby photo of another person (e.g., a child, a spouse, etc.), the smart assistant device automatically places a call to the son. Such configuration enables users to place calls without having to remember telephone numbers, or other user identifiers of other persons.

The total number of symbols or types of input that can be provided in AAC channels is virtually inexhaustible, e.g., different types of detectable facial expression (examples: eye movement, eyebrow movement, lip movement etc.); body movements or language (examples: waving of an arm, walking past the smart assistant, etc.); body-part movement (examples: blowing of air through the mouth, curling the fingers of a hand); movements of physical objects; sounds created by physical objects (examples: a set of bells of differing pitches); etc., can each be associated with unique commands.

The human-machine interaction channels, per techniques of this disclosure, have at least the following features.

- They are accessible to people with different abilities.
- They are simple to use by an end-user and simple to configure, e.g., by a caregiver.
- They are personalized, e.g., can be adapted to the unique skills or capabilities of the user.
- They create a logged-in experience for the user while obviating a login procedure.
- They couple physical action by the user to commands for the smart assistant device.

AAC communication with a smart assistant device or other machine, per the techniques described herein, is performed only with the permission of the user and caregiver. Such permission is modifiable and can be revoked at any time.

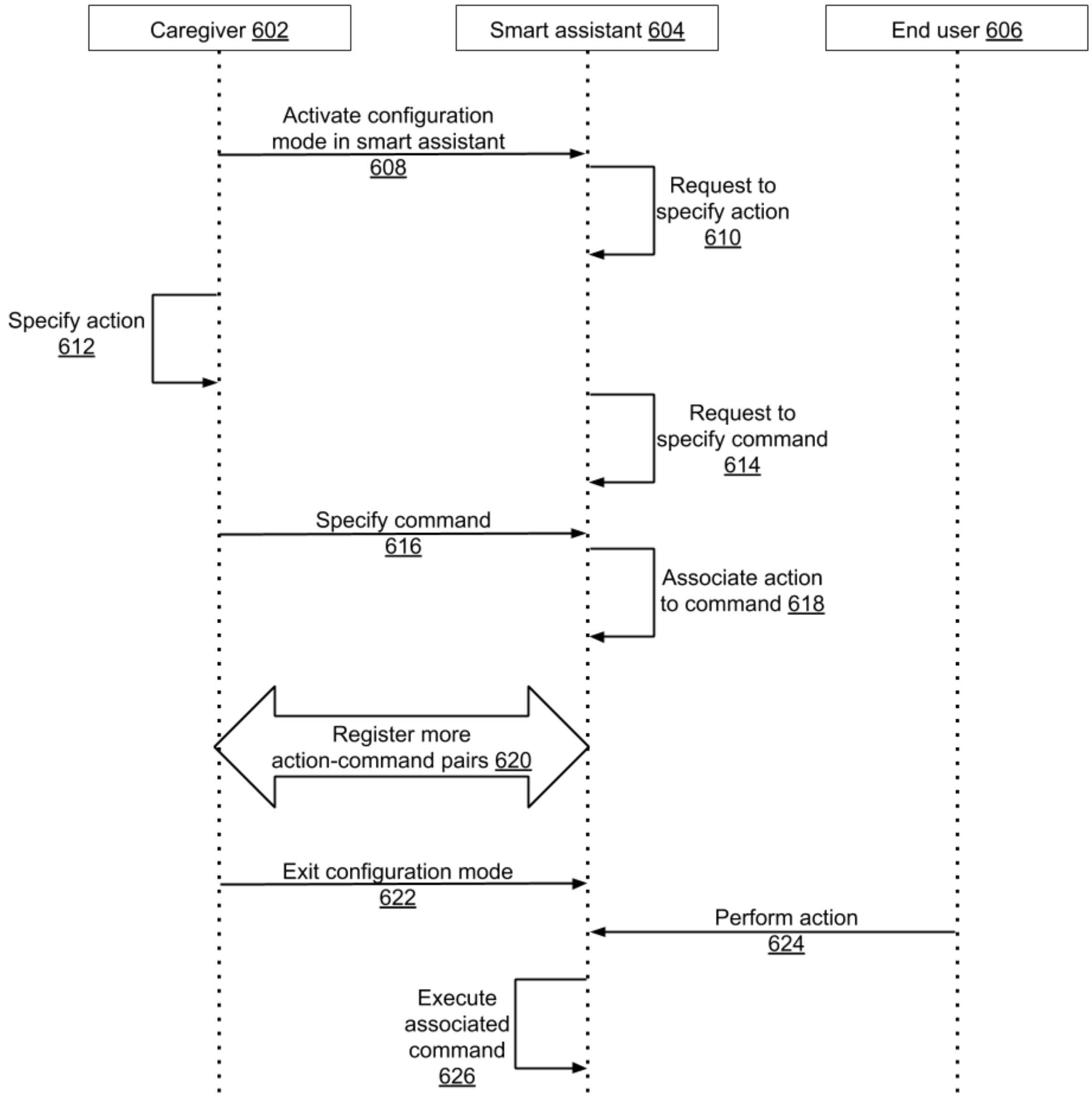

Fig. 6: Configuring and using a smart assistant under an AAC framework

Fig. 6 illustrates configuring of a smart assistant with action-command pairs and usage of a configured smart assistant by an end-user. The configuration may be for a particular smart assistant application that executes on one or more smart assistant devices such as smartphones, tablets, speakers, wearable devices, home appliances, etc.

A caregiver (602) activates smart assistant (604) in configuration mode (608). Such activation can be performed by touching or pressing a pre-configured button, by long-pressing a pre-configured button, by voice-commanding the smart assistant to assert configuration mode, etc. In response to a command to start the configuration mode, the smart assistant requests the caregiver to specify an action (610). The caregiver specifies the action (612), e.g., by performing the action. As explained above, example actions can be moving particular physical objects, particular body or body-part movements, gestures, facial expressions, etc. The smart assistant recognizes the action and issues a request to the caregiver to specify a command (614) to be associated with the just-performed action. The caregiver specifies the command (616). As explained above, example commands can include operating an IoT device, playing particular pieces of media or launching media channels, performing an Internet search and reporting results, the asking of a question, etc. The smart assistant associates the recognized action to command (618) and stores the association as configuration information.

This action-command sequence can be repeated as often as necessary such that one or more action-command pairs are registered with the smart assistant (620). Any one action is at most associated with one command, and vice-versa. When a sufficient number of action-command pairs are registered, the user issues a command to the smart assistant to exit configuration mode (622). After the configuration, when it is detected that an end-user has performed a certain registered action (624), the smart assistant executes the command associated with that action (626).

The techniques of the disclosure can be implemented in smart assistant software applications and/or dedicated hardware. Still alternately, dedicated hardware can be built as an accessibility gateway to a smart assistant and to the smart assistant ecosystem, e.g., IoT devices, internet, etc.

By expanding the possibilities for human-machine interaction, the techniques described herein open up the benefits of consumer technology to those that may have been excluded or had limited access to such technology, e.g., differently-abled people, senior citizens, etc. Upon implementation with consent of a parent and/or caregiver, the techniques can also serve as a learning tool for young children, e.g., by associating one action with one command, the principle of cause and effect is demonstrated. Smart assistant actions can be easily configured, e.g., by a caregiver, and provide a simple physical user interface. The techniques create a logged-in experience (e.g., being able to download or play music from an online music service) without requiring the user to perform login actions. The techniques of this disclosure can also be generalized to engage the end-user in a game, e.g., the completion of a puzzle, quiz, or other challenge. In this manner, the techniques enable individuals such as senior citizens, people with motor disabilities, etc. to access consumer technology devices and services without excessive reliance on others and provide a degree of autonomy.

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

Smart assistants utilize speech recognition, sensing, artificial intelligence, and networking technologies to enable improved human-machine interaction. However, there remain use cases where smart assistants are not easily usable by humans. For example, voice-activated assistants are not accessible to those that are hearing or speech impaired. Touchscreen based assistants are not usable by those who lack fine motor skills and/or reading ability. While caregivers can assist differently-abled individuals with consumer technology, this sometimes can have the effect of reinforcing in the differently-abled individual their lack of autonomy. This disclosure adds to the modalities or channels by which humans can control and communicate with smart assistants by enabling use of physical objects, facial expressions, gross motor skills, body movements, etc. to provide commands. Collectively, these techniques of control and communication are referred to as alternative and augmentative communication (AAC).