US 20200103963A1

(19) United States

(12) Patent Application Publication

KELLY et al.

(10) Pub. No.: US 2020/0103963 A1

(43) Pub. Date: Apr. 2, 2020

(54) **DEVICE CONTROL USING GAZE INFORMATION**

(71) Applicant: **Apple Inc.**, Cupertino, CA (US)

(72) Inventors: **Sean B. KELLY**, San Francisco, CA (US); **Felipe BACIM DE ARAUJO E SILVA**, Santa Clara, CA (US); **Karlin Y. BARK**, Menlo Park, CA (US)

(21) Appl. No.: **16/553,622**

(22) Filed: **Aug. 28, 2019**

Related U.S. Application Data

(60) Provisional application No. 62/739,087, filed on Sep. 28, 2018.

Publication Classification

(51) **Int. Cl.**

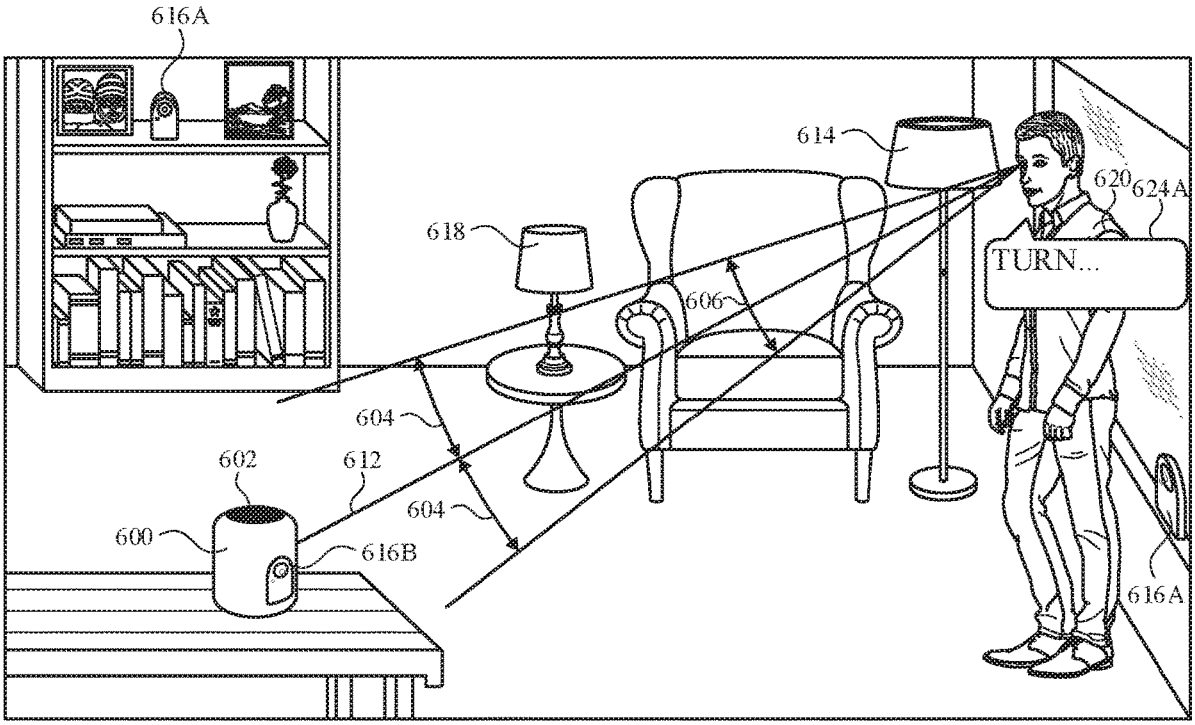
G06F 3/01	(2006.01)
G10L 15/22	(2006.01)
G06F 3/16	(2006.01)

(52) **U.S. Cl.**

CPC ***G06F 3/013*** (2013.01); ***G10L 15/22*** (2013.01); *G10L 2015/223* (2013.01); ***G06F 3/167*** (2013.01); ***G06F 3/165*** (2013.01)

(57) **ABSTRACT**

The present disclosure generally relates to controlling electronic devices. In some examples, the electronic device uses gaze information to activate a digital assistant. In some examples, the electronic device uses gaze information to identify an external device on which to act. In some examples, the electronic device provides an indication that distinguishes between different speakers.



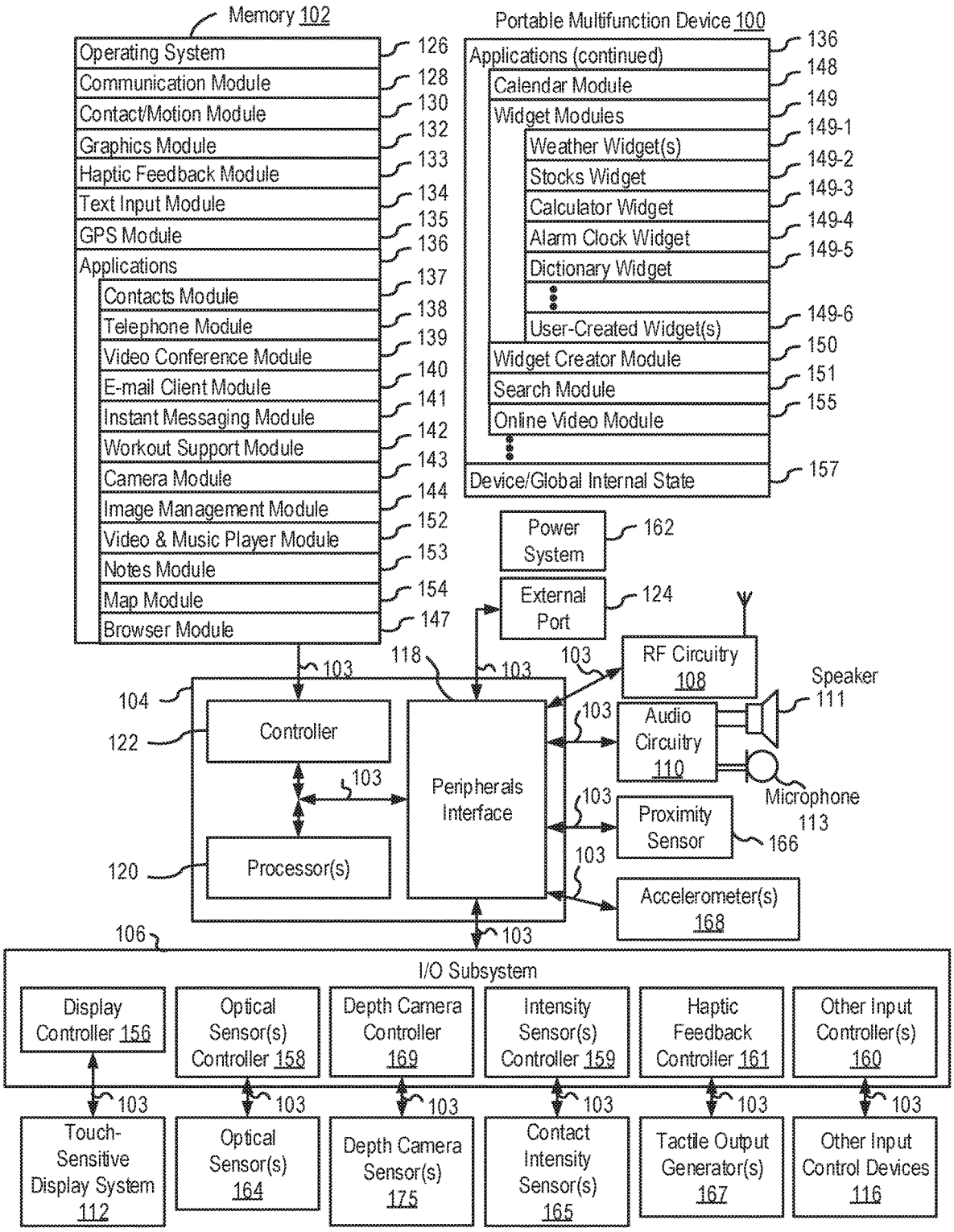

Memory 102
Operating System 126
Communication Module 128
Contact/Motion Module 130
Graphics Module 132
Haptic Feedback Module 133
Text Input Module 134
GPS Module 135
Applications 136
Contacts Module 137
Telephone Module 138
Video Conference Module 139
E-mail Client Module 140
Instant Messaging Module 141
Workout Support Module 142
Camera Module 143
Image Management Module 144
Video & Music Player Module 152
Notes Module 153
Map Module 154
Browser Module 147
Portable Multifunction Device 100
Applications (continued) 136
Calendar Module 148
Widget Modules 149
Weather Widget(s) 149-1
Stocks Widget 149-2
Calculator Widget 149-3
Alarm Clock Widget 149-4
Dictionary Widget 149-5
User-Created Widget(s) 149-6
Widget Creator Module 150
Search Module 151
Online Video Module 155
Device/Global Internal State 157
Power System 162
External Port 124
RF Circuitry 108
Speaker 111
Audio Circuitry 110
Microphone 113
Proximity Sensor 166
Accelerometer(s) 168
Controller 122
Processor(s) 120
Peripherals Interface 118
103
104
106
I/O Subsystem
Display Controller 156
Optical Sensor(s) Controller 158
Depth Camera Controller 169
Intensity Sensor(s) Controller 159
Haptic Feedback Controller 161
Other Input Controller(s) 160
Touch-Sensitive Display System 112
Optical Sensor(s) 164
Depth Camera Sensor(s) 175
Contact Intensity Sensor(s) 165
Tactile Output Generator(s) 167
Other Input Control Devices 116

FIG. 1A

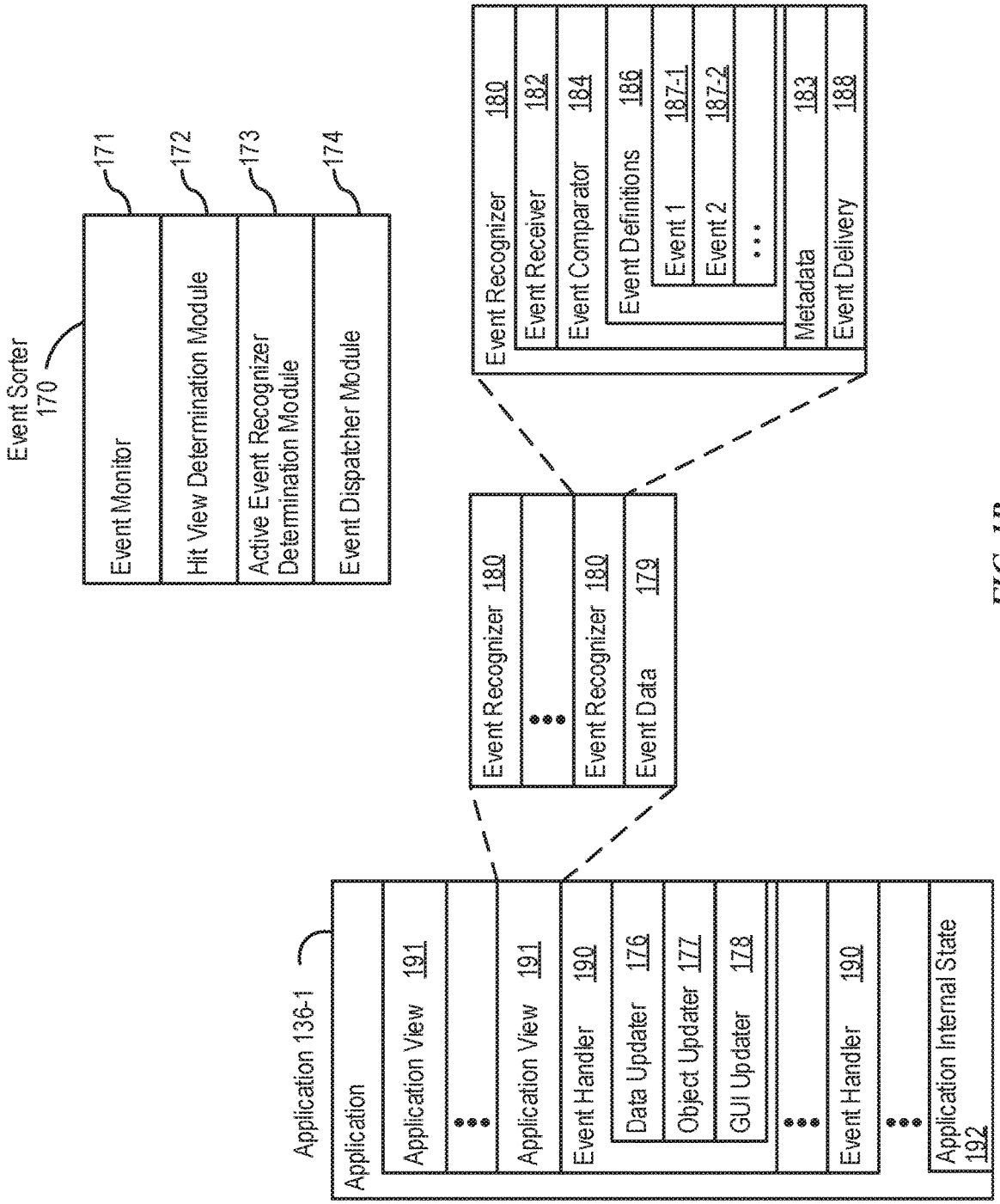

Event Sorter
170
171
Event Monitor
172
Hit View Determination Module
173
Active Event Recognizer Determination Module
174
Event Dispatcher Module
Event Recognizer 180
Event Receiver 182
Event Comparator 184
Event Definitions 186
Event 1 187-1
Event 2 187-2
...
Metadata 183
Event Delivery 188
Event Recognizer 180
...
Event Recognizer 180
Event Data 179
Application 136-1
Application
Application View 191
...
Application View 191
Event Handler 190
Data Updater 176
Object Updater 177
GUI Updater 178
...
Event Handler 190
...
Application Internal State 192

FIG. 1B

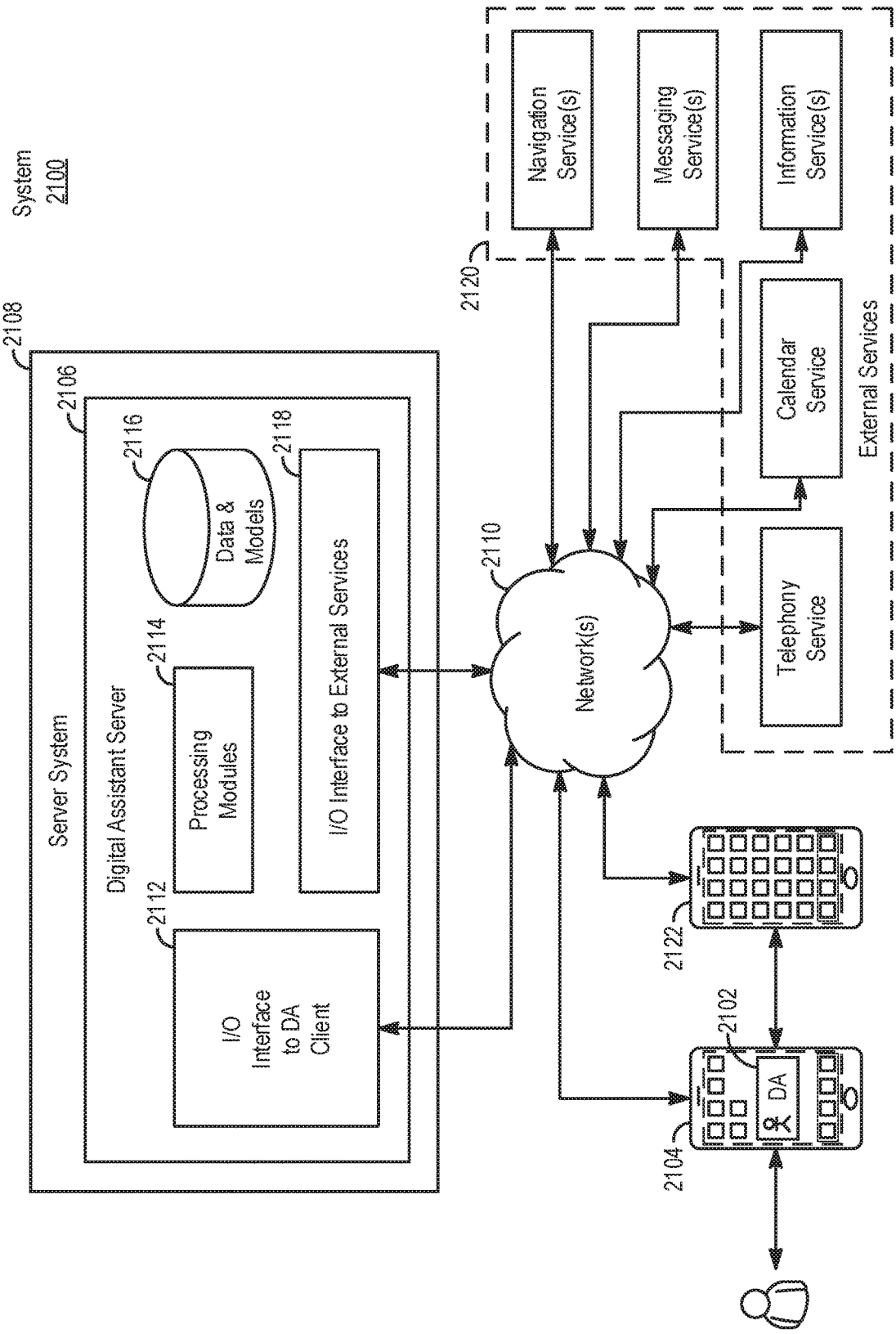

System
2100
2108
Server System
2106
Digital Assistant Server
2116
Data & Models
2118
I/O Interface to External Services
2114
Processing Modules
2112
I/O Interface to DA Client
2110
Network(s)
2120
Navigation Service(s)
Messaging Service(s)
Information Service(s)
Calendar Service
Telephony Service
External Services
2122
2102
DA
2104

FIG. 1C

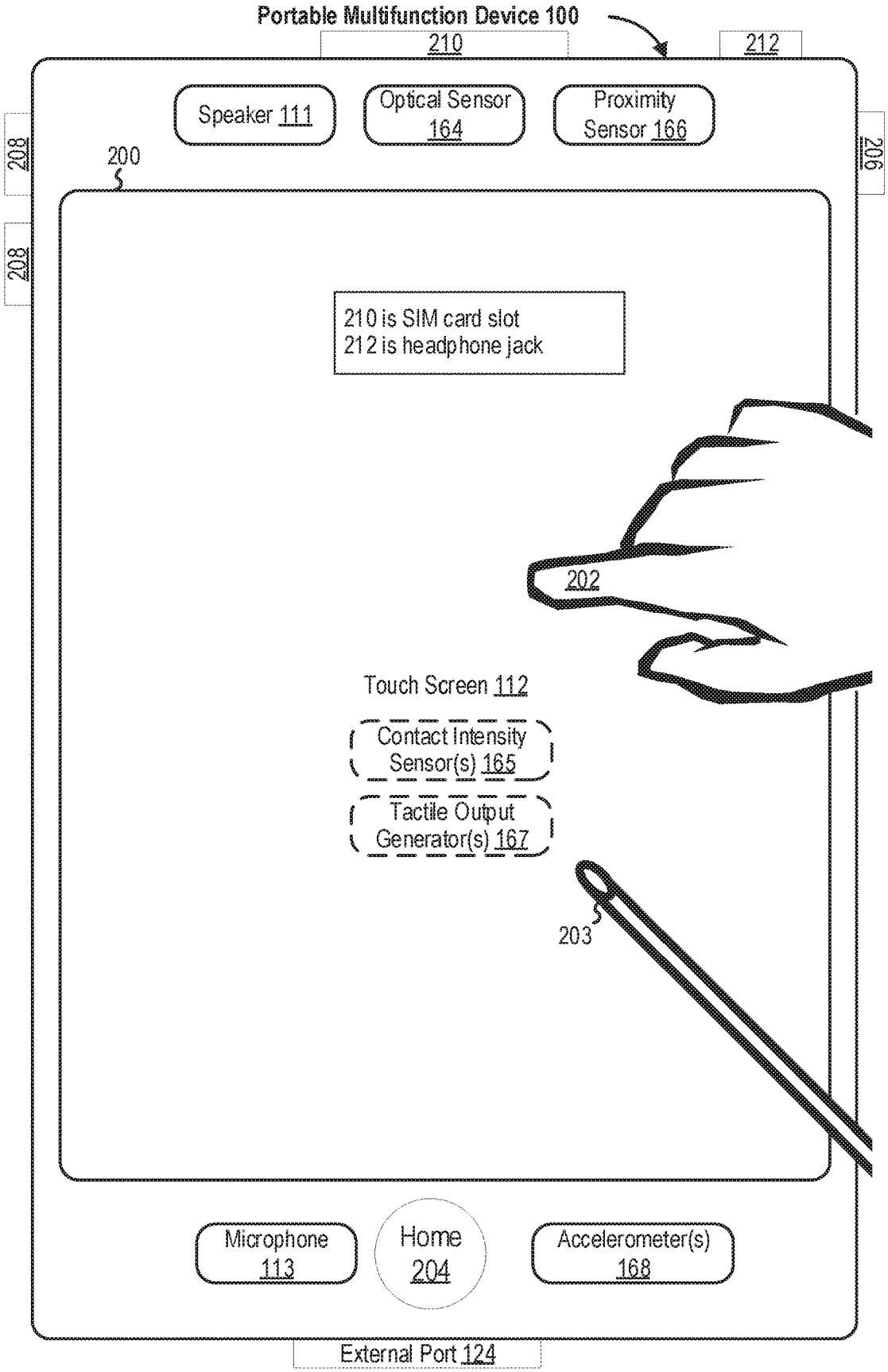

Portable Multifunction Device 100
210
212
Speaker 111
Optical Sensor 164
Proximity Sensor 166
208
206
200
208
210 is SIM card slot
212 is headphone jack
202
Touch Screen 112
Contact Intensity Sensor(s) 165
Tactile Output Generator(s) 167
203
Microphone 113
Home 204
Accelerometer(s) 168
External Port 124

FIG. 2

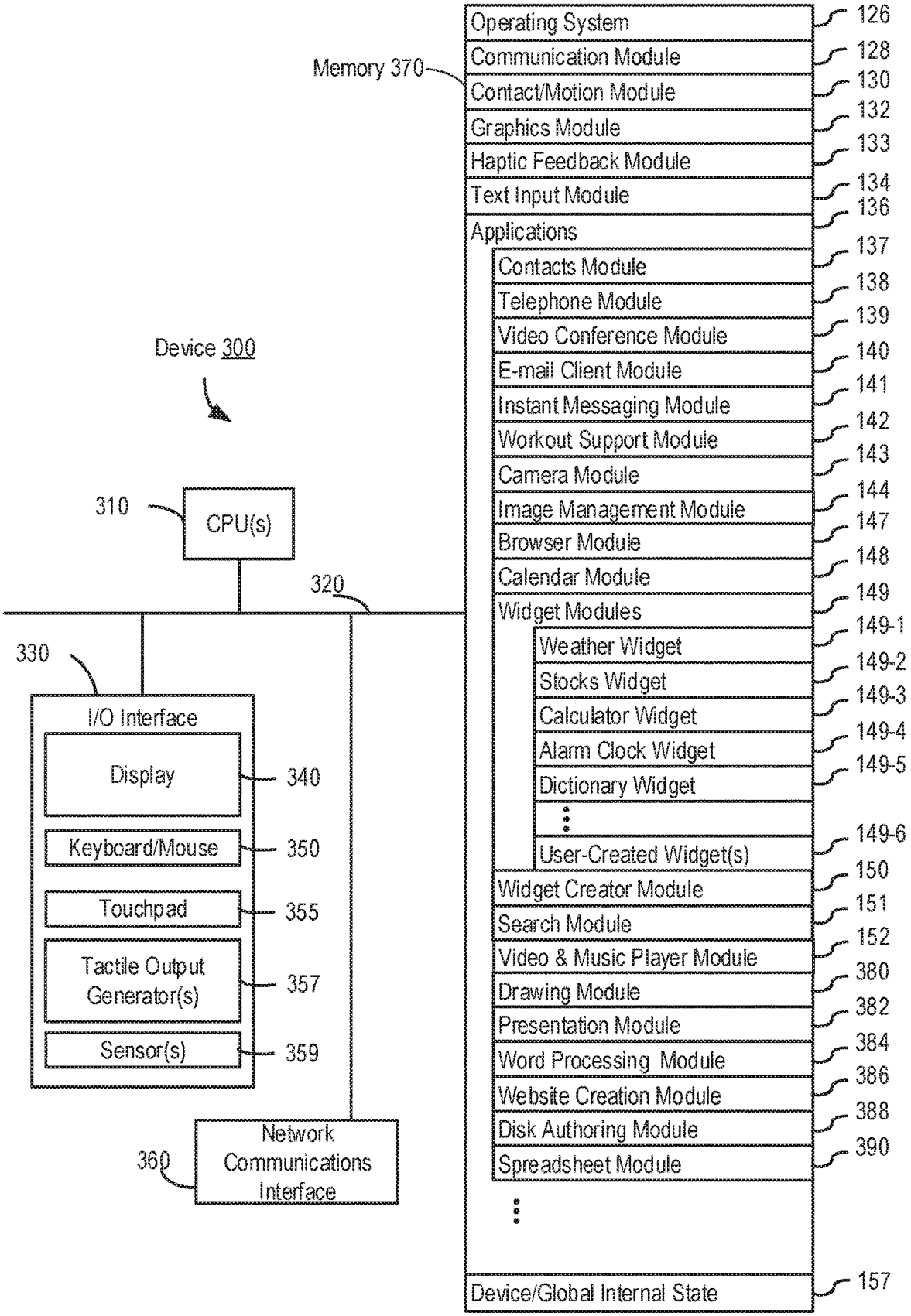

Device 300
310
CPU(s)
320
330
I/O Interface
Display
340
Keyboard/Mouse
350
Touchpad
355
Tactile Output Generator(s)
357
Sensor(s)
359
Network Communications Interface
360
Memory 370
Operating System
126
Communication Module
128
Contact/Motion Module
130
Graphics Module
132
Haptic Feedback Module
133
Text Input Module
134
Applications
136
Contacts Module
137
Telephone Module
138
Video Conference Module
139
E-mail Client Module
140
Instant Messaging Module
141
Workout Support Module
142
Camera Module
143
Image Management Module
144
Browser Module
147
Calendar Module
148
Widget Modules
149
Weather Widget
149-1
Stocks Widget
149-2
Calculator Widget
149-3
Alarm Clock Widget
149-4
Dictionary Widget
149-5
User-Created Widget(s)
149-6
Widget Creator Module
150
Search Module
151
Video & Music Player Module
152
Drawing Module
380
Presentation Module
382
Word Processing Module
384
Website Creation Module
386
Disk Authoring Module
388
Spreadsheet Module
390
Device/Global Internal State
157

FIG. 3

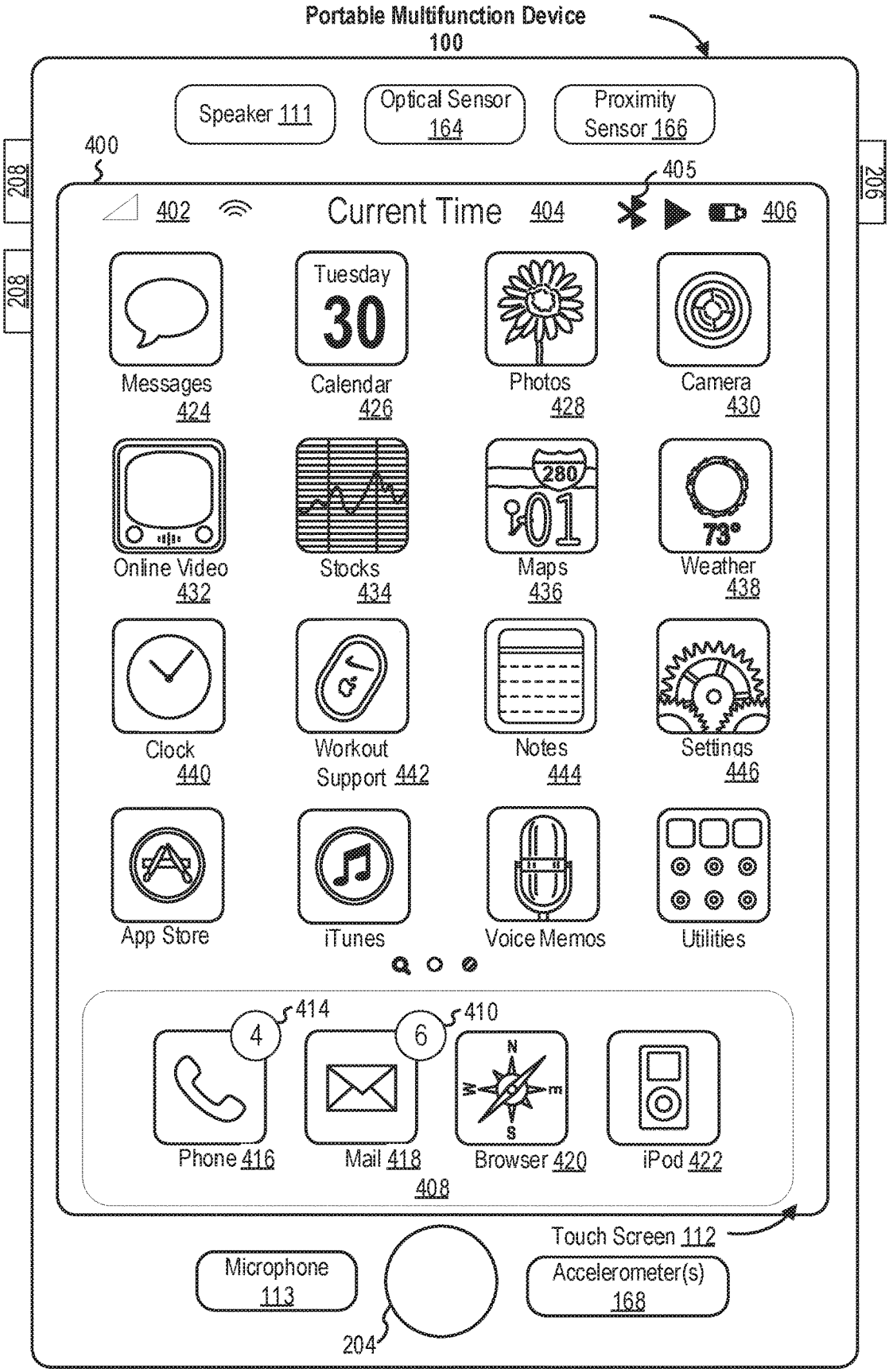
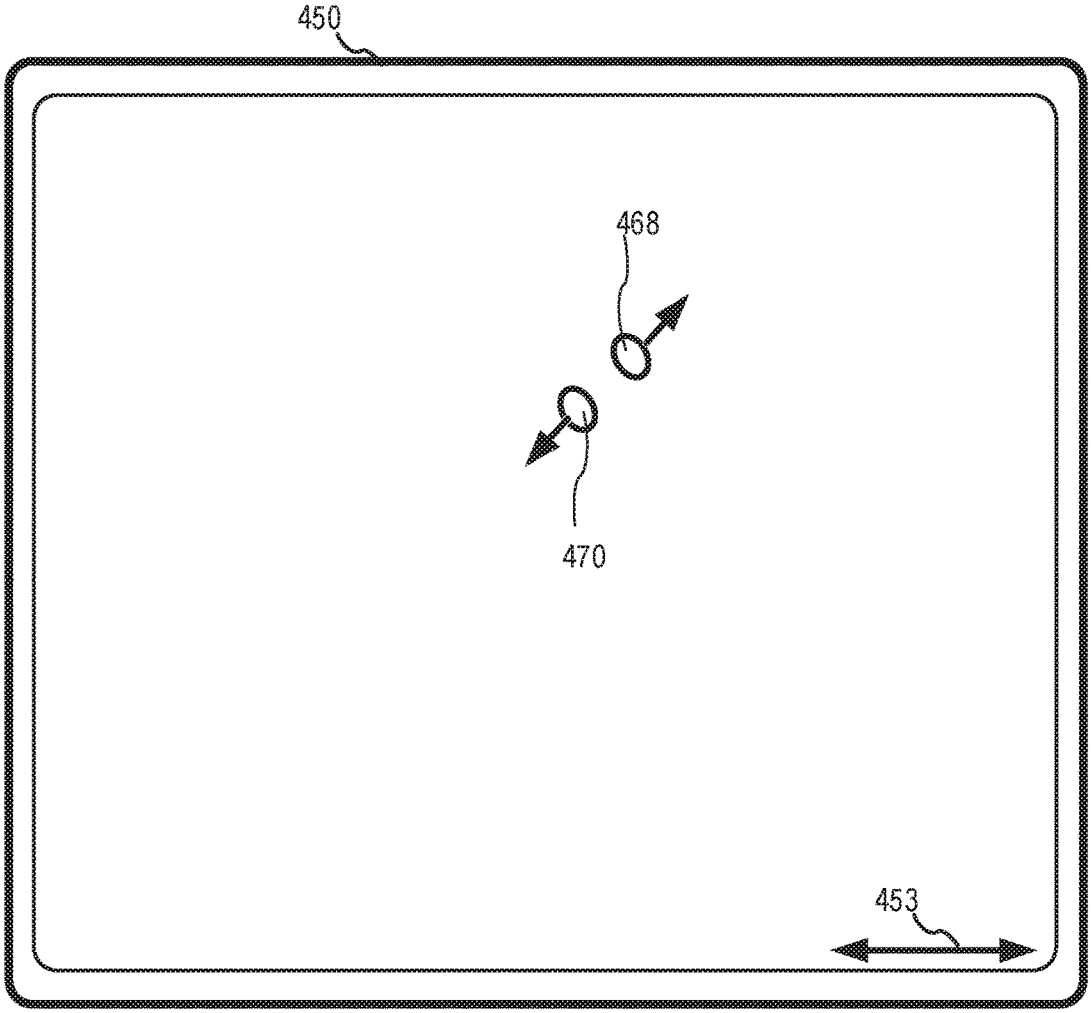

Portable Multifunction Device
100
Speaker 111
Optical Sensor 164
Proximity Sensor 166
400
405
208
208
206
402
Current Time 404
406
Tuesday
30
Messages 424
Calendar 426
Photos 428
Camera 430
Online Video 432
Stocks 434
280
101
Maps 436
73°
Weather 438
Clock 440
Workout Support 442
Notes 444
Settings 446
App Store
iTunes
Voice Memos
Utilities
414
410
4
6
N
W
E
S
Phone 416
Mail 418
Browser 420
iPod 422
408
Touch Screen 112
Microphone 113
Accelerometer(s) 168
204

FIG. 4A


450
468
470
453

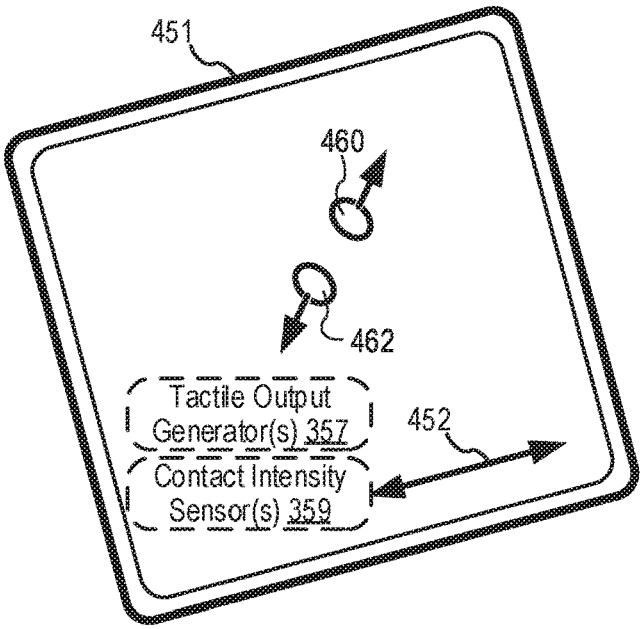

451
460
462
Tactile Output Generator(s) 357
452
Contact Intensity Sensor(s) 359

FIG. 4B

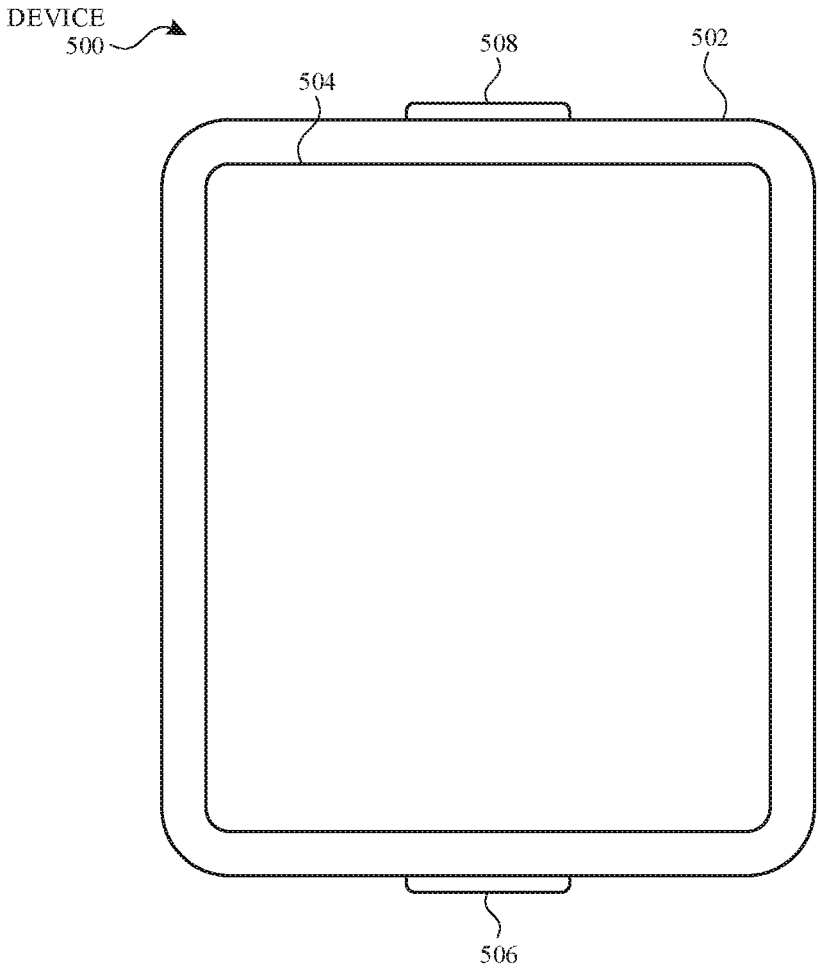


FIG. 5A

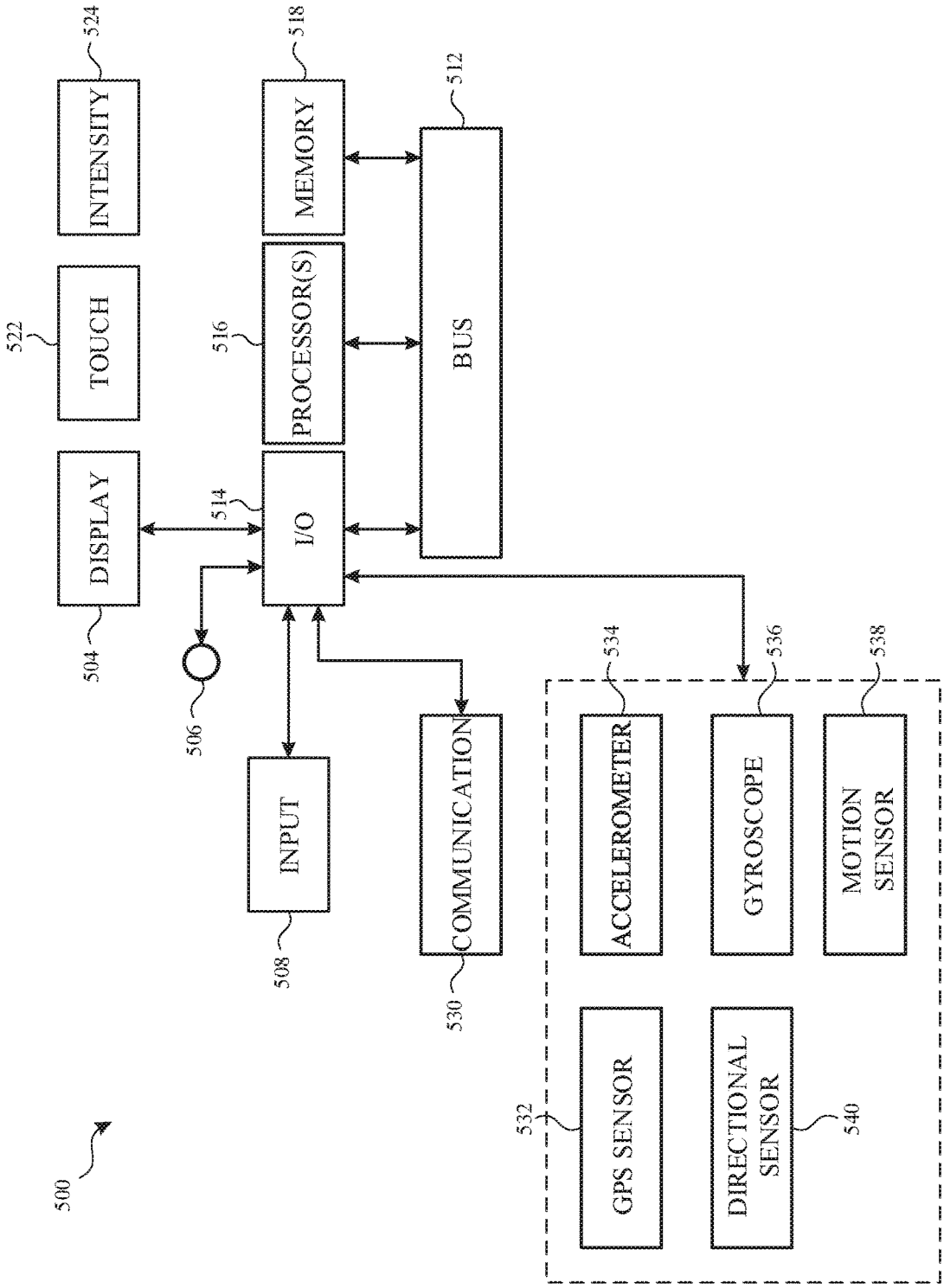

500
DISPLAY
504
TOUCH
522
INTENSITY
524
506
I/O
514
PROCESSOR(S)
516
MEMORY
518
BUS
512
INPUT
508
COMMUNICATION
530
GPS SENSOR
532
ACCELEROMETER
534
DIRECTIONAL SENSOR
540
GYROSCOPE
536
MOTION SENSOR
538

FIG. 5B

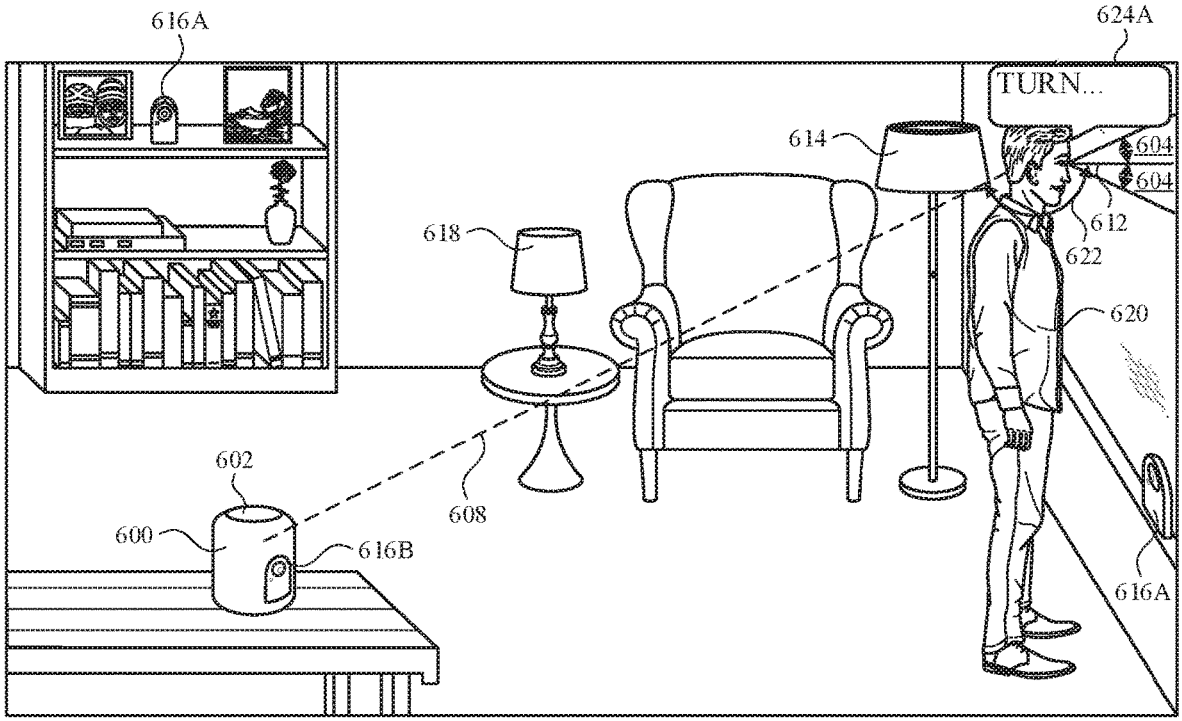

616A
624A
TURN...
614
604
604
612
622
618
620
602
608
600
616B
616A

FIG. 6A

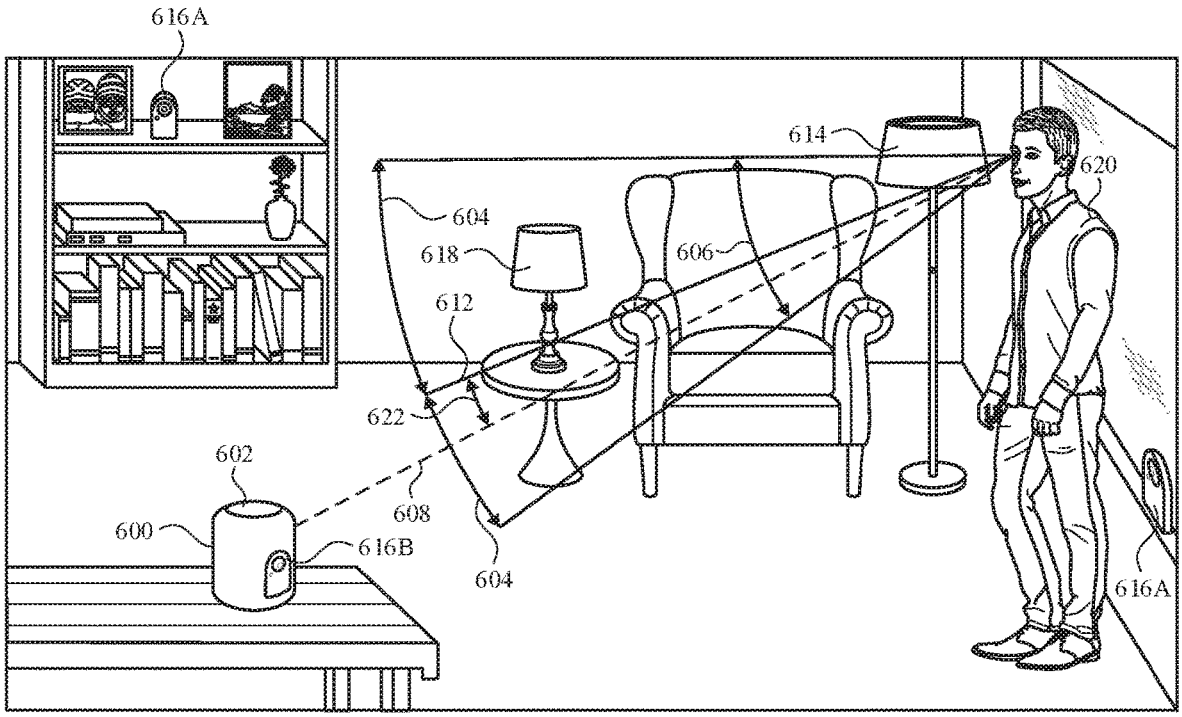

616A
614
620
604
618
606
612
622
602
608
600
616B
604
616A

FIG. 6B

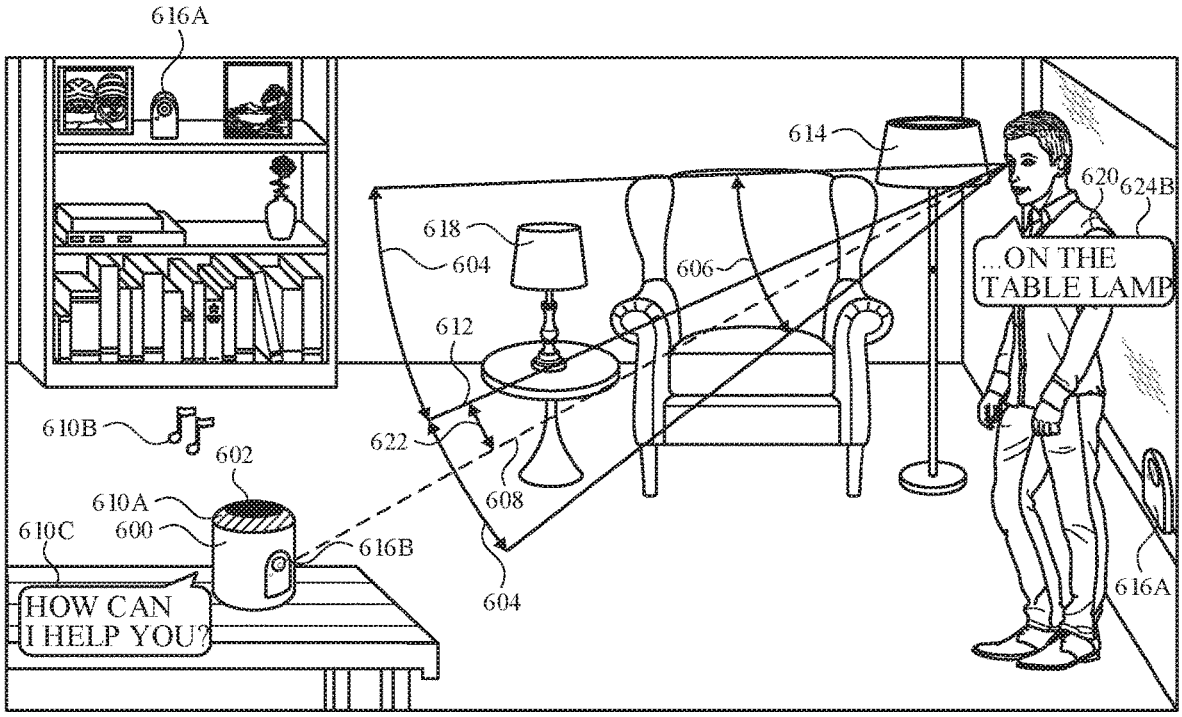

616A
614
620
624B
618
604
606
...ON THE
TABLE LAMP
612
610B
602
622
610A
608
610C
600
616B
604
616A
HOW CAN
I HELP YOU?

FIG. 6C

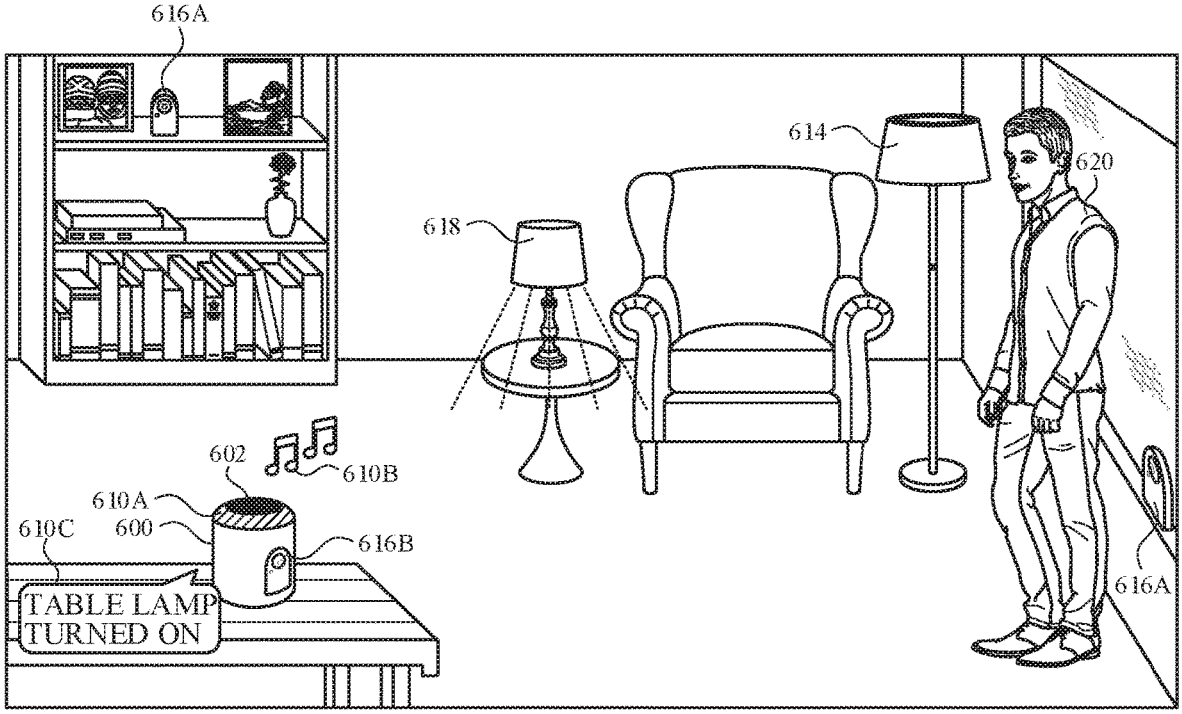

616A
614
620
618
602
610B
610A
610C
600
616B
616A
TABLE LAMP
TURNED ON

FIG. 6D

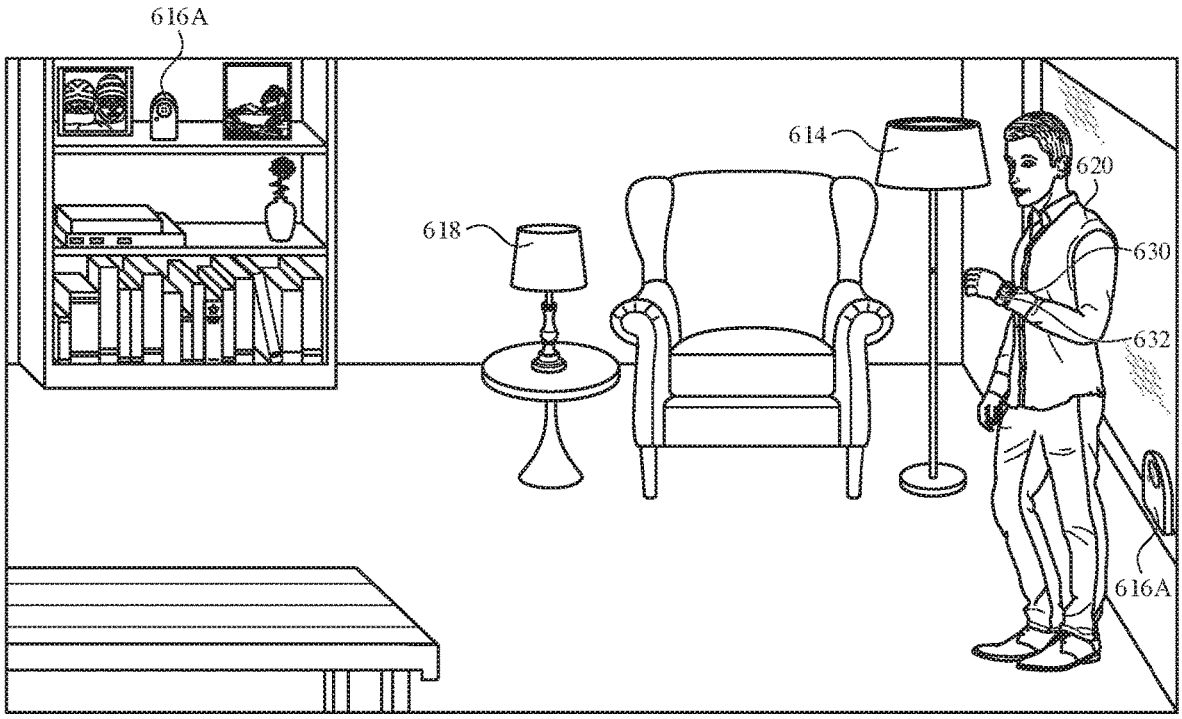

616A
614
620
618
630
632
616A

FIG. 6E

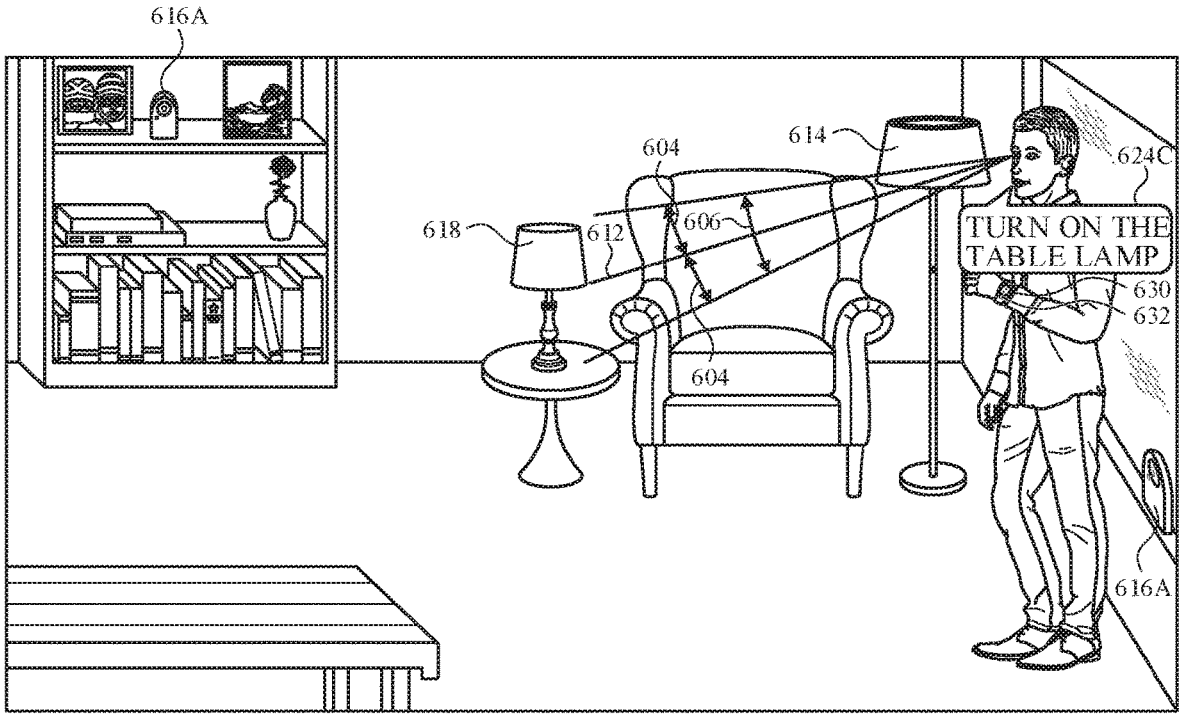

616A
604
614
624C
618
612
606
TURN ON THE
TABLE LAMP
630
632
604
616A

FIG. 6F

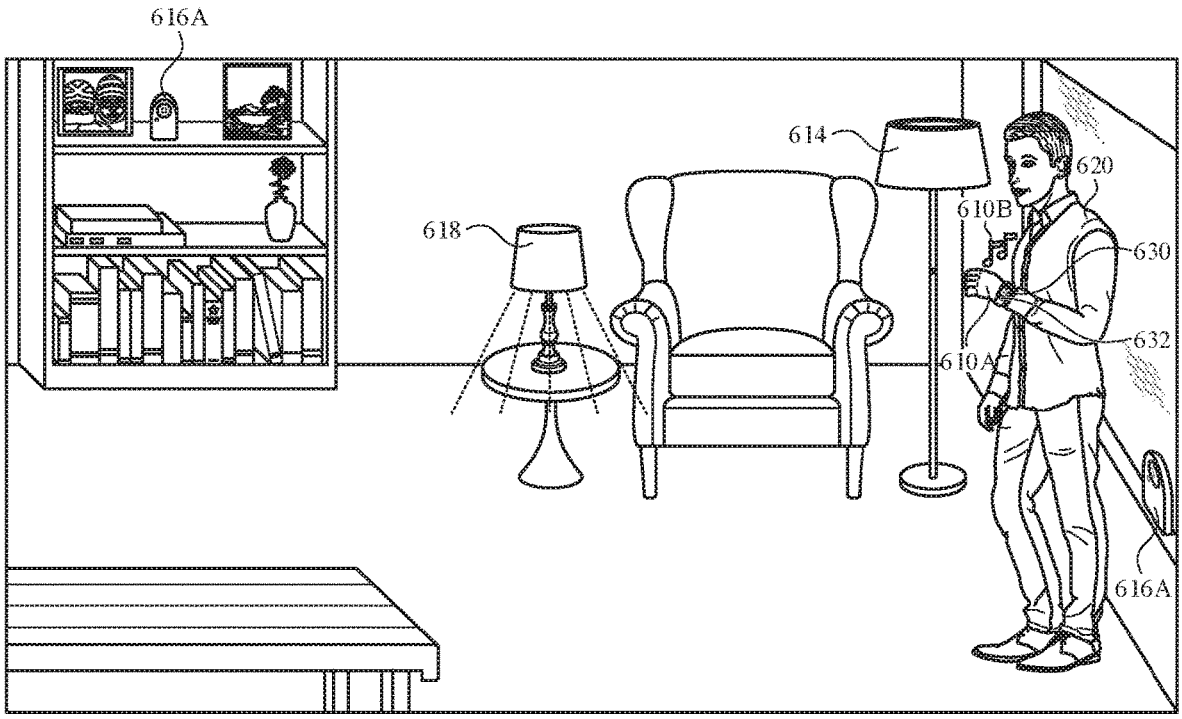

616A
614
620
610B
618
630
632
610A
616A

FIG. 6G

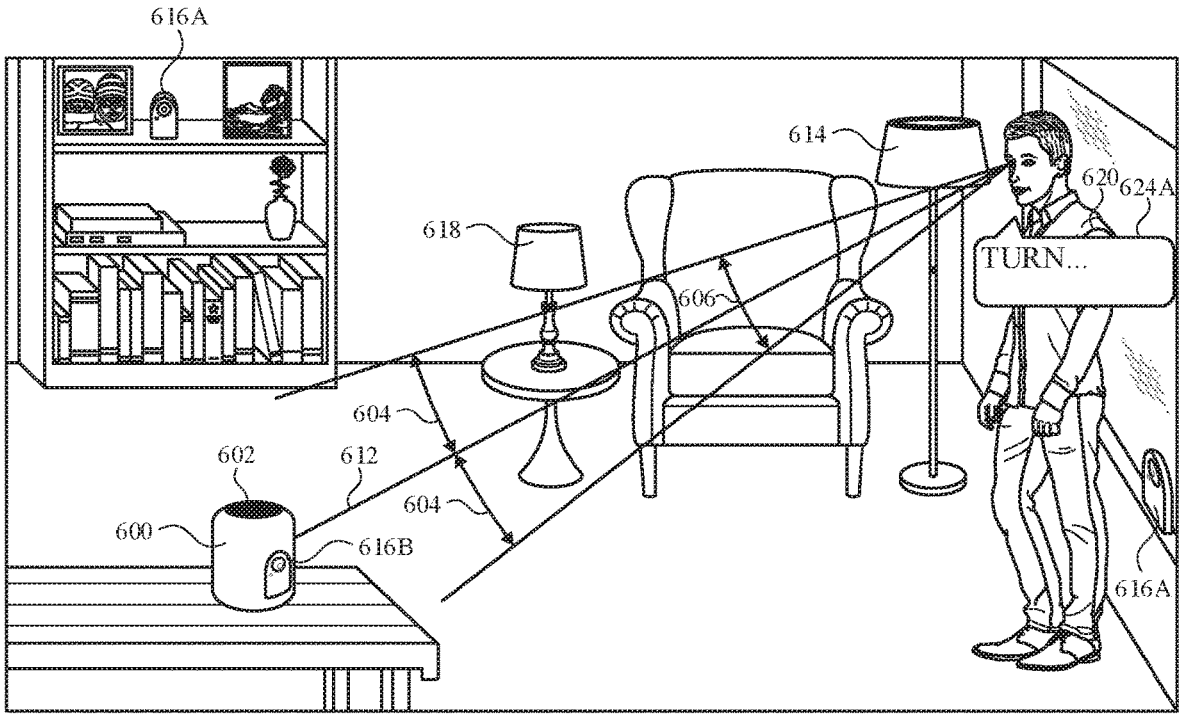

616A
614
620
624A
618
TURN...
606
604
602
612
604
600
616B
616A

FIG. 6H

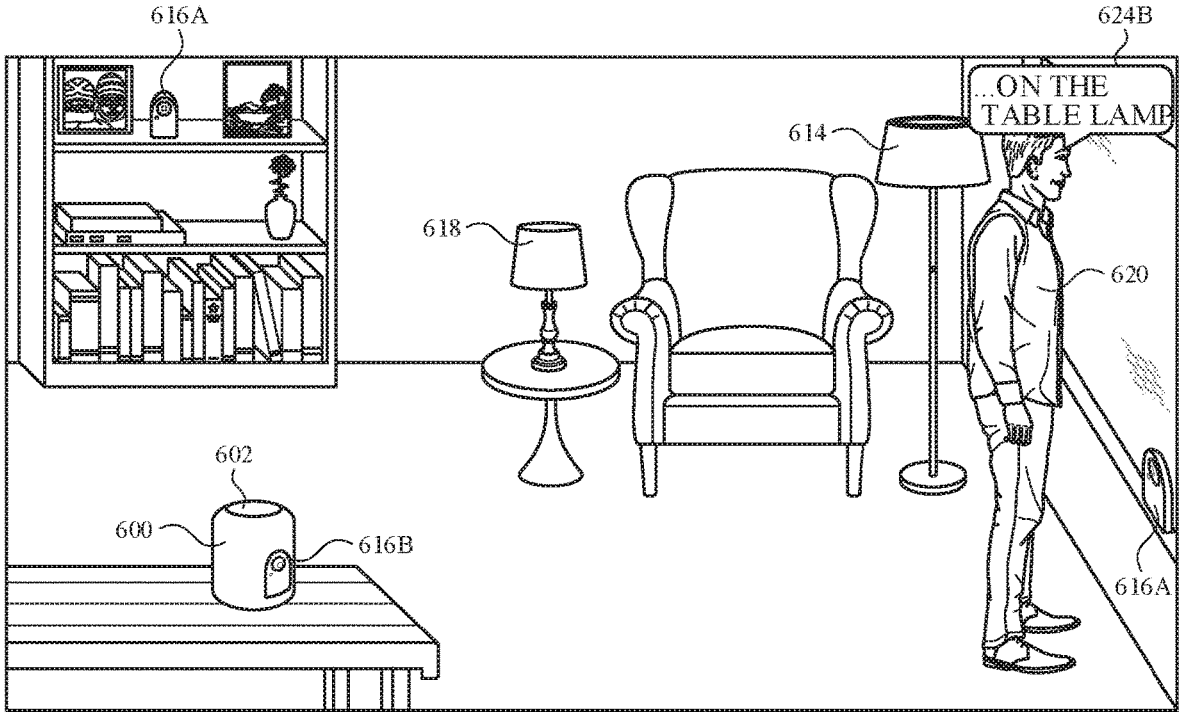

616A
624B
...ON THE
TABLE LAMP
614
618
620
602
600
616B
616A

FIG. 6I

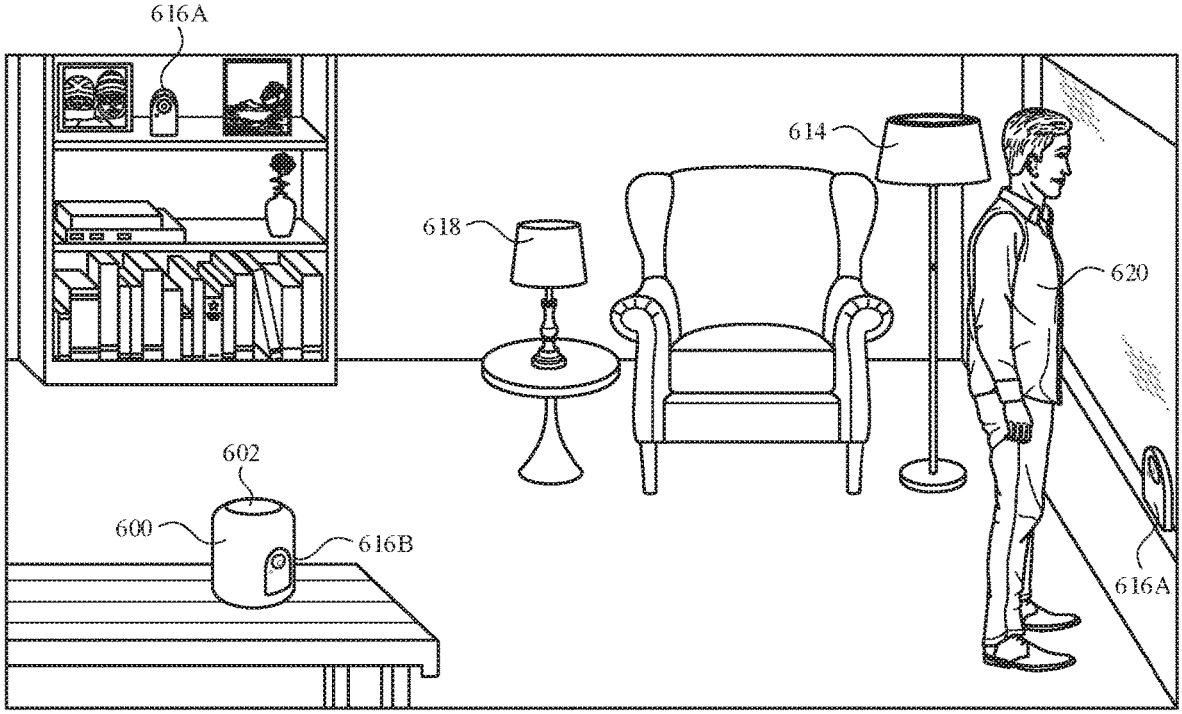

616A
614
618
620
602
600
616B
616A

FIG. 6J

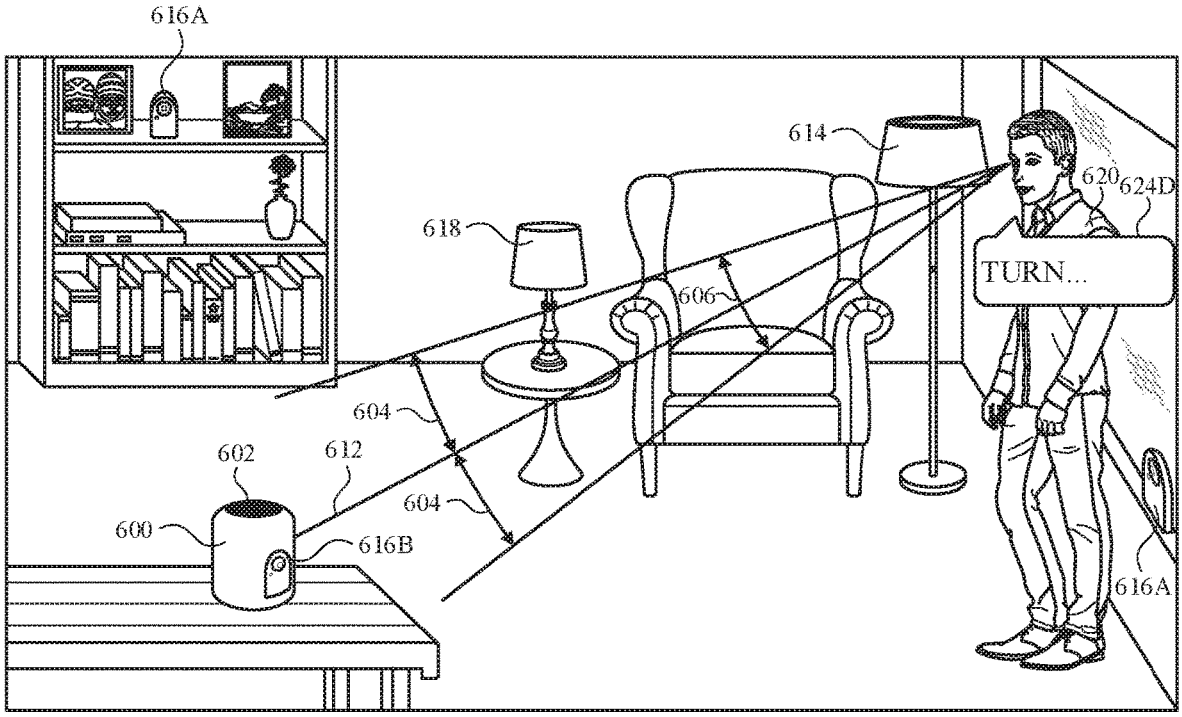

616A
614
620
624D
618
TURN...
606
604
602
612
604
600
616B
616A

FIG. 6K

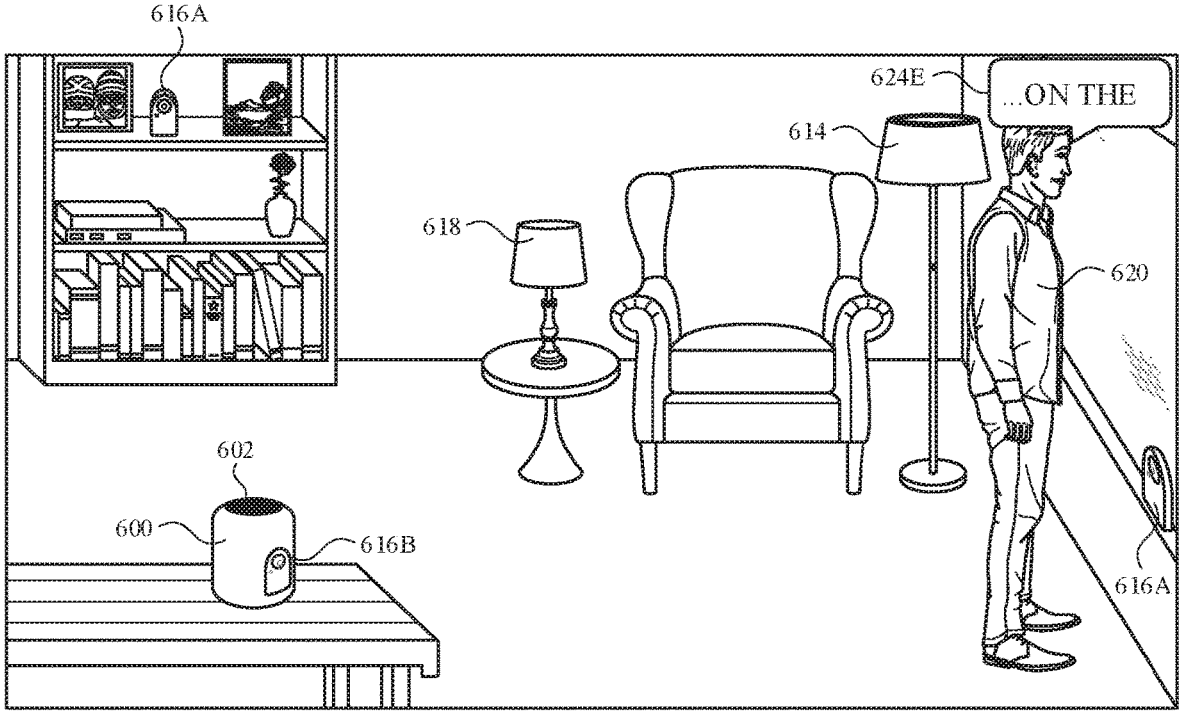

616A
624E
...ON THE
614
618
620
602
600
616B
616A

FIG. 6L

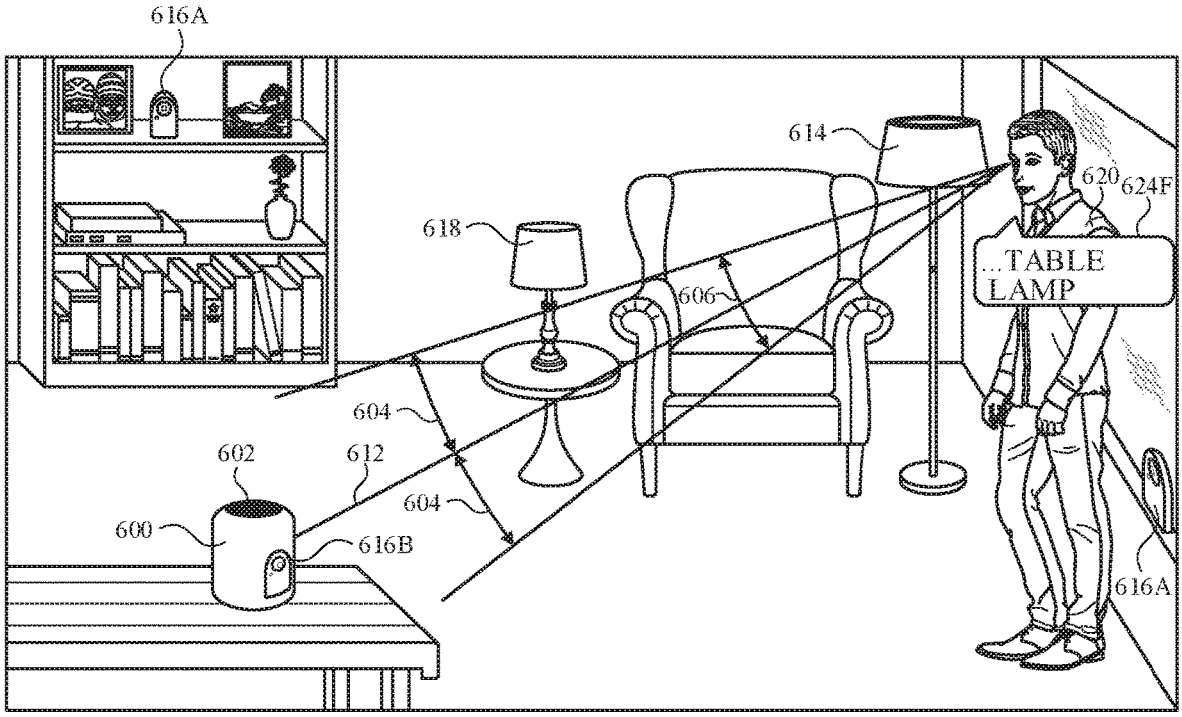

616A
614
620
624F
618
...TABLE LAMP
606
604
602
612
604
600
616B
616A

FIG. 6M

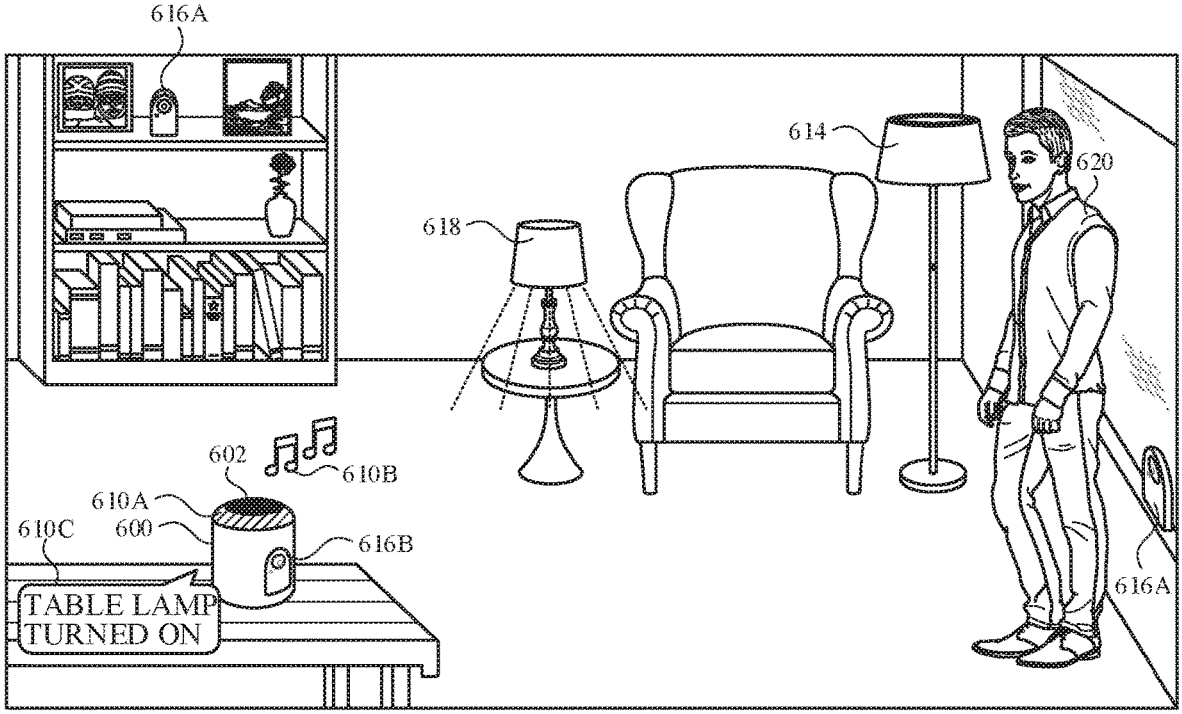

616A
614
620
618
602
610B
610A
610C
600
616B
TABLE LAMP TURNED ON
616A

FIG. 6N

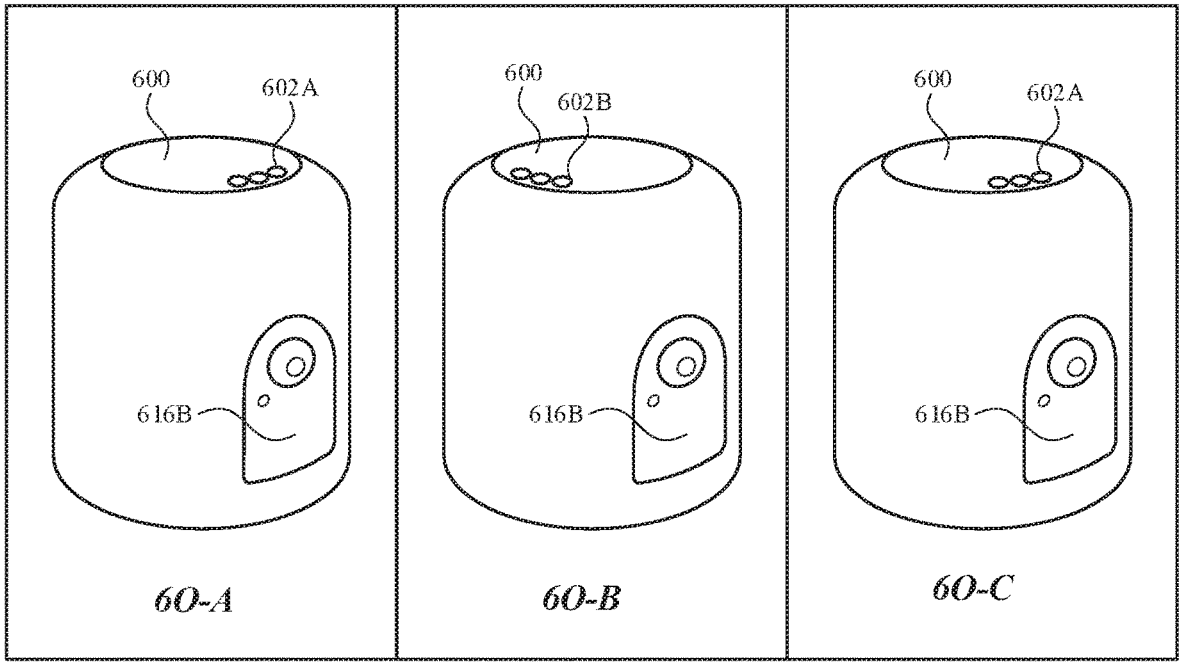


FIG. 6O

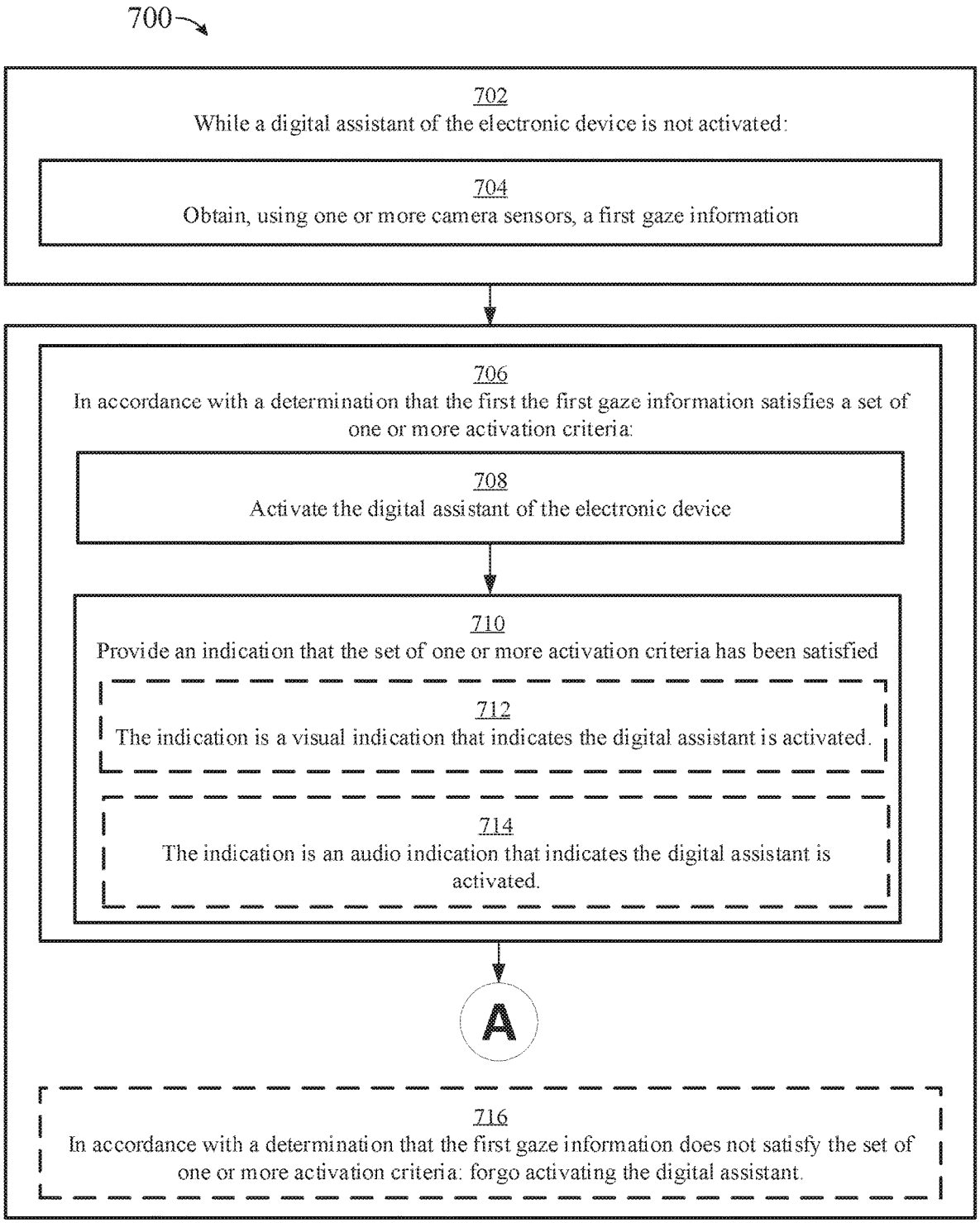

700
702
While a digital assistant of the electronic device is not activated:
704
Obtain, using one or more camera sensors, a first gaze information
706
In accordance with a determination that the first the first gaze information satisfies a set of one or more activation criteria:
708
Activate the digital assistant of the electronic device
710
Provide an indication that the set of one or more activation criteria has been satisfied
712
The indication is a visual indication that indicates the digital assistant is activated.
714
The indication is an audio indication that indicates the digital assistant is activated.
A
716
In accordance with a determination that the first gaze information does not satisfy the set of one or more activation criteria: forgo activating the digital assistant.

FIG. 7A

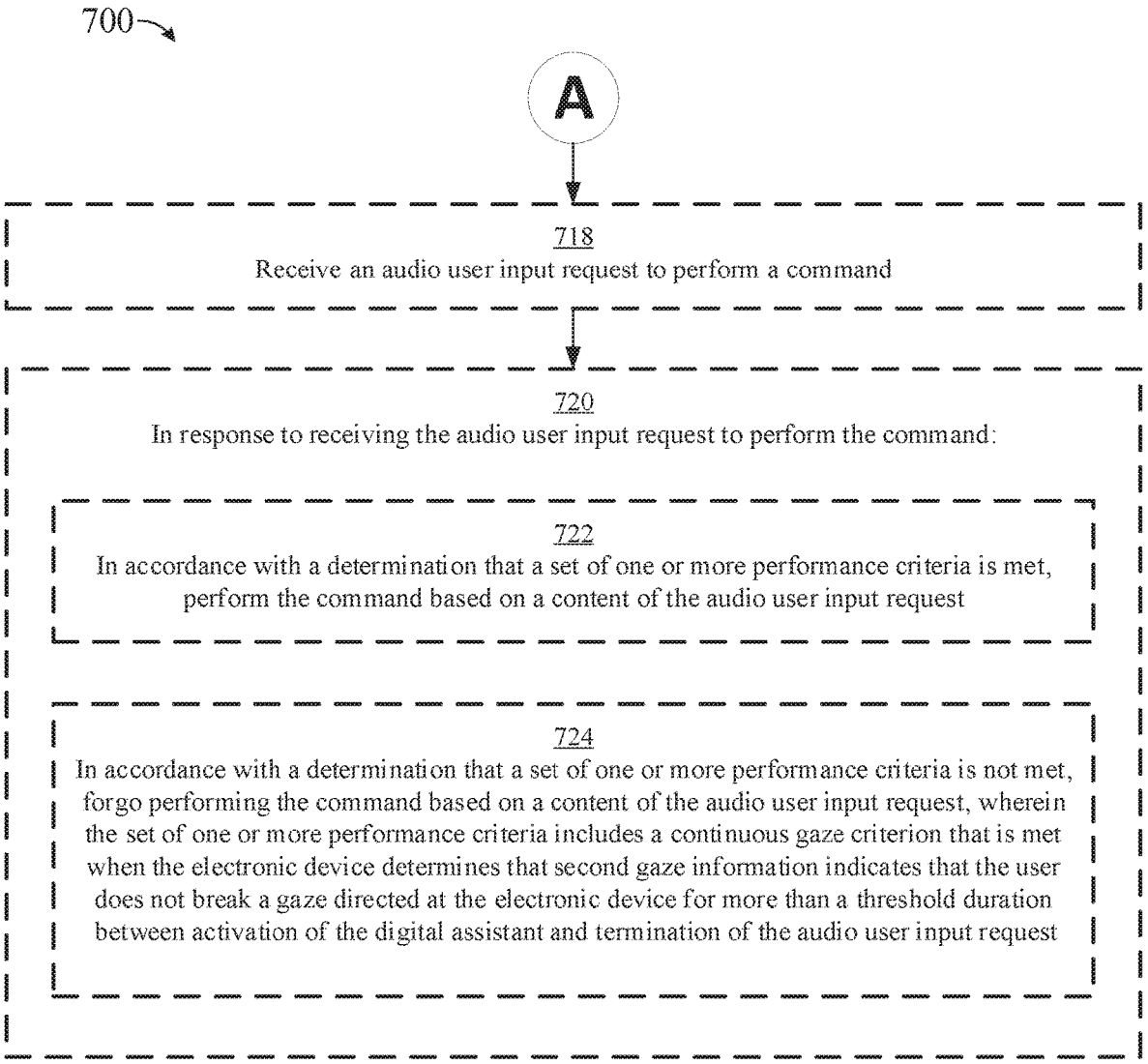

700
A
718
Receive an audio user input request to perform a command
720
In response to receiving the audio user input request to perform the command:
722
In accordance with a determination that a set of one or more performance criteria is met, perform the command based on a content of the audio user input request
724
In accordance with a determination that a set of one or more performance criteria is not met, forgo performing the command based on a content of the audio user input request, wherein the set of one or more performance criteria includes a continuous gaze criterion that is met when the electronic device determines that second gaze information indicates that the user does not break a gaze directed at the electronic device for more than a threshold duration between activation of the digital assistant and termination of the audio user input request

FIG. 7B

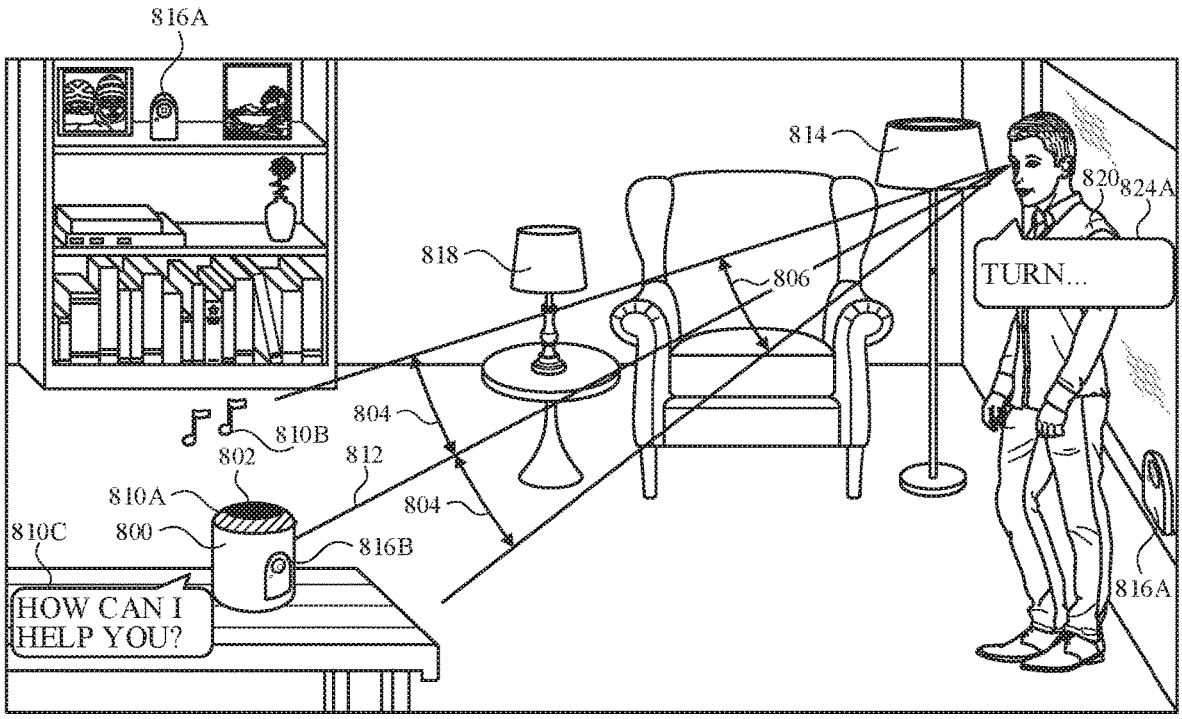

816A
814
820
824A
818
806
TURN...
804
810B
802
812
810A
804
810C
800
816B
HOW CAN I HELP YOU?
816A

FIG. 8A

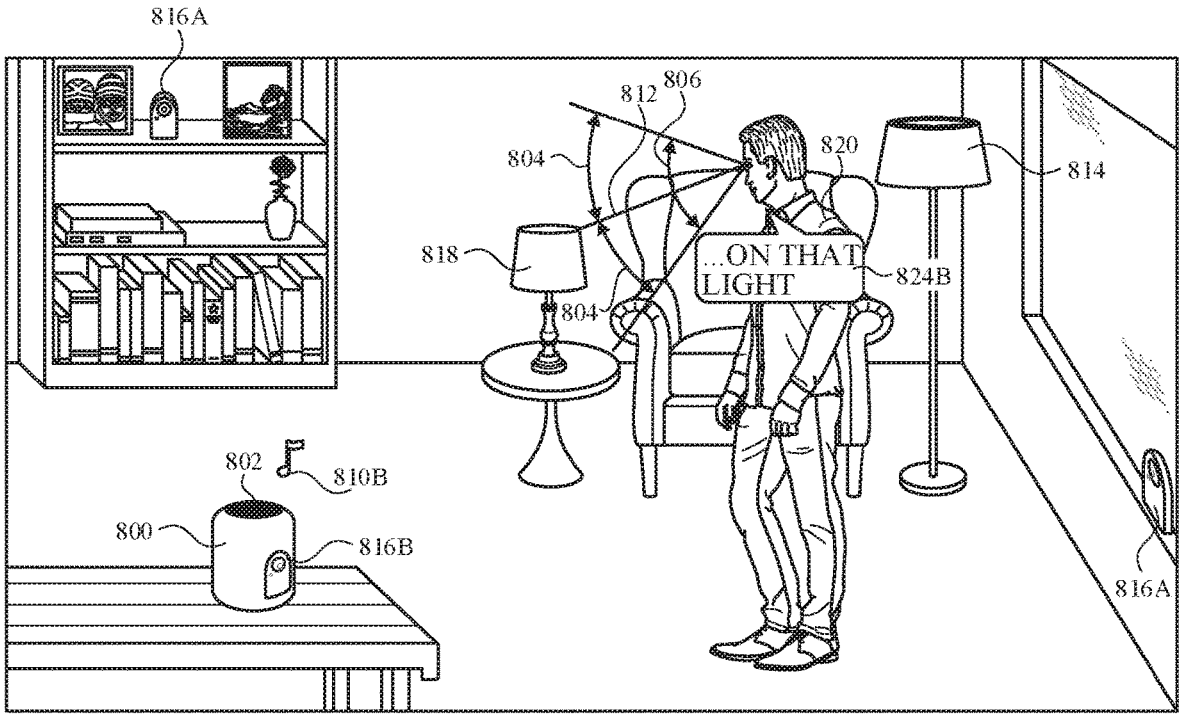

816A
812
806
804
820
814
818
...ON THAT LIGHT
824B
804
802
810B
800
816B
816A

FIG. 8B

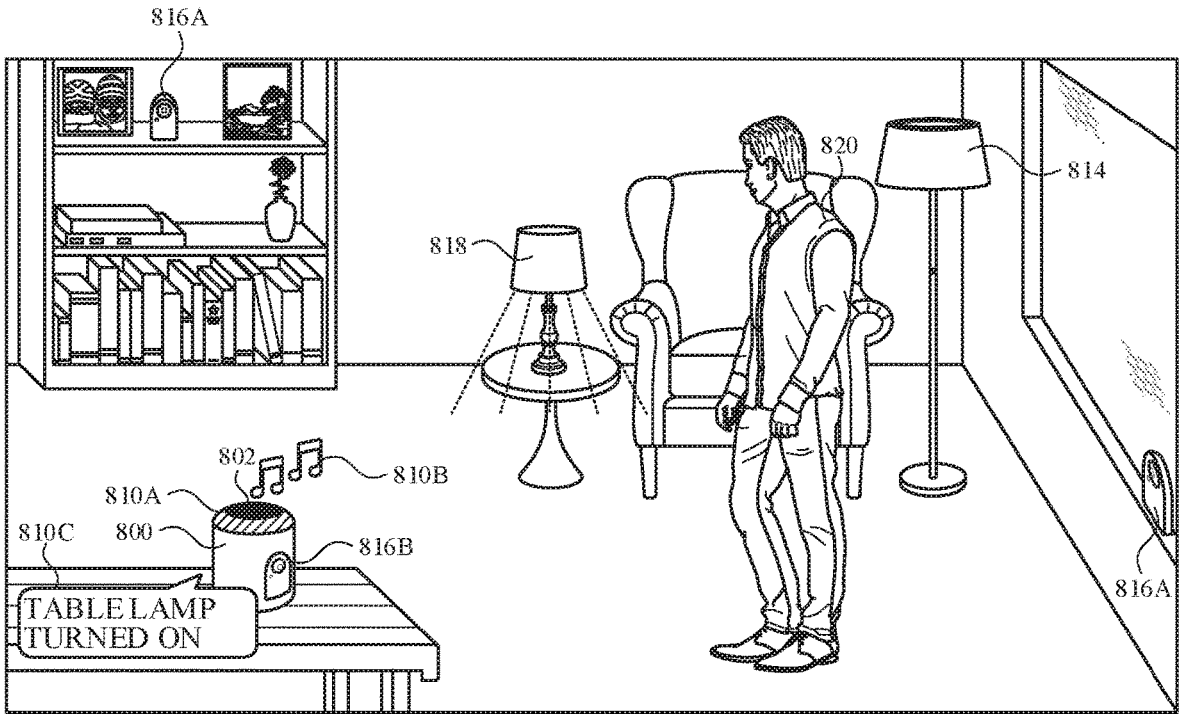

816A
820
814
818
802
810B
810A
810C
800
816B
TABLE LAMP
TURNED ON
816A

FIG. 8C

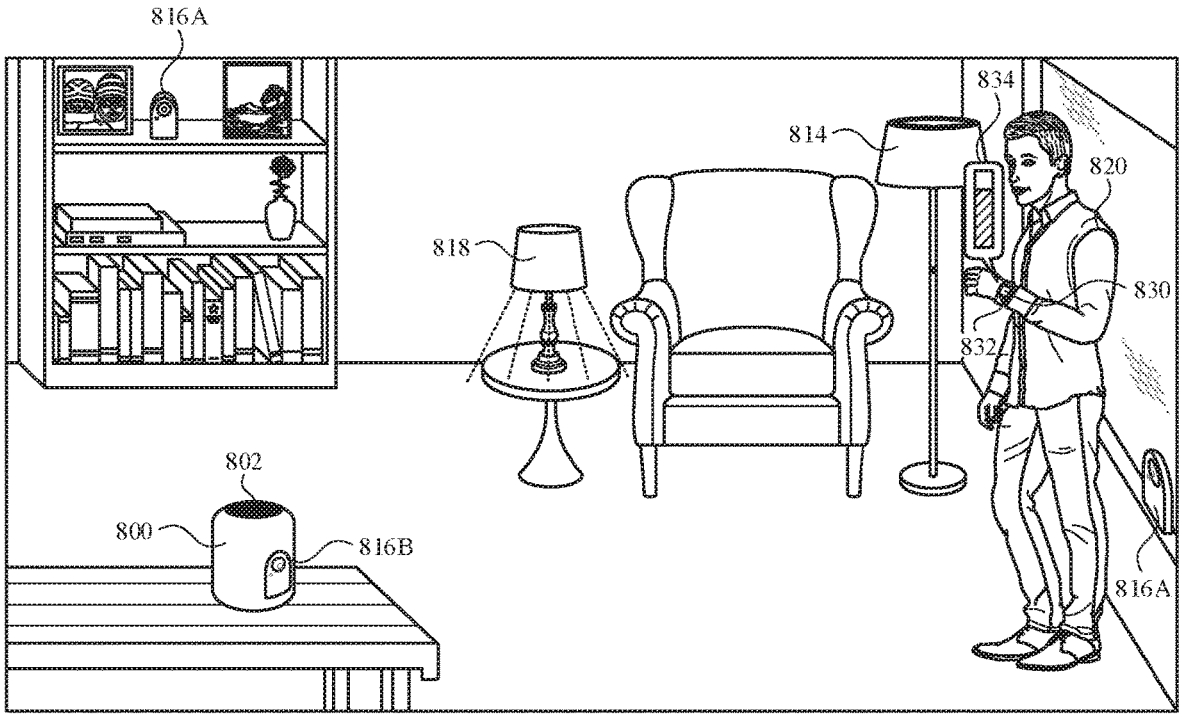

816A
834
814
820
818
830
832
802
800
816B
816A

FIG. 8D

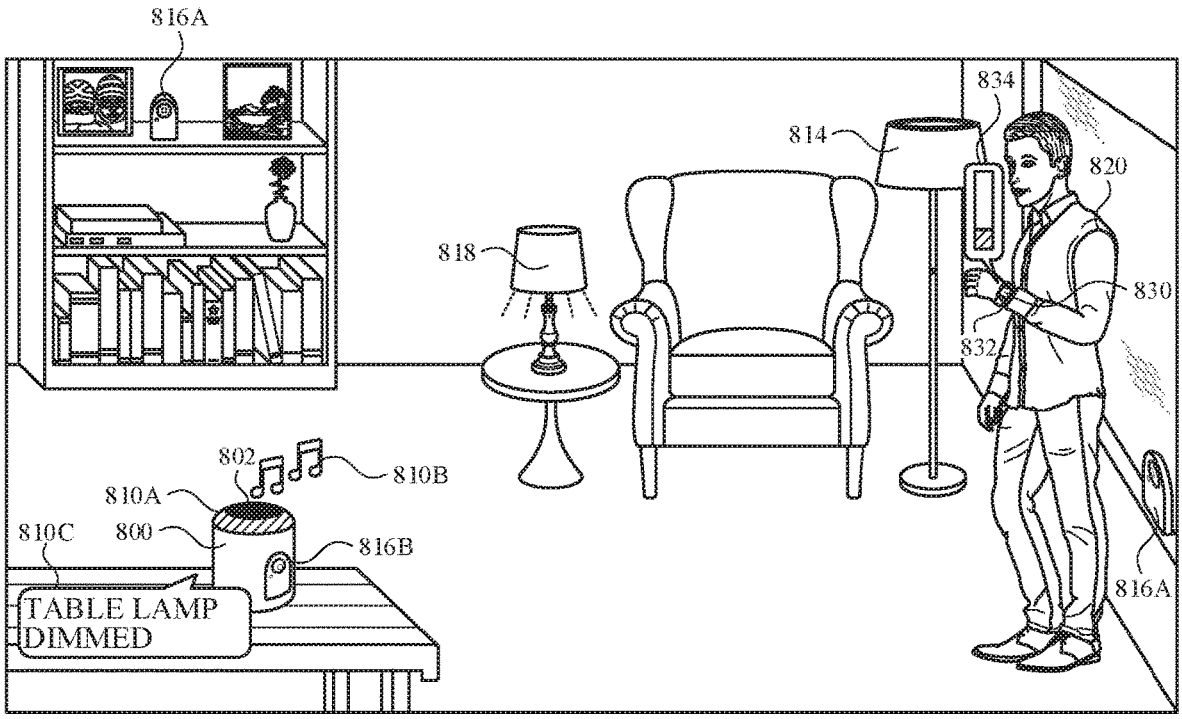

816A
834
814
820
818
830
832
802
810B
810A
810C
800
816B
TABLE LAMP DIMMED
816A

FIG. 8E

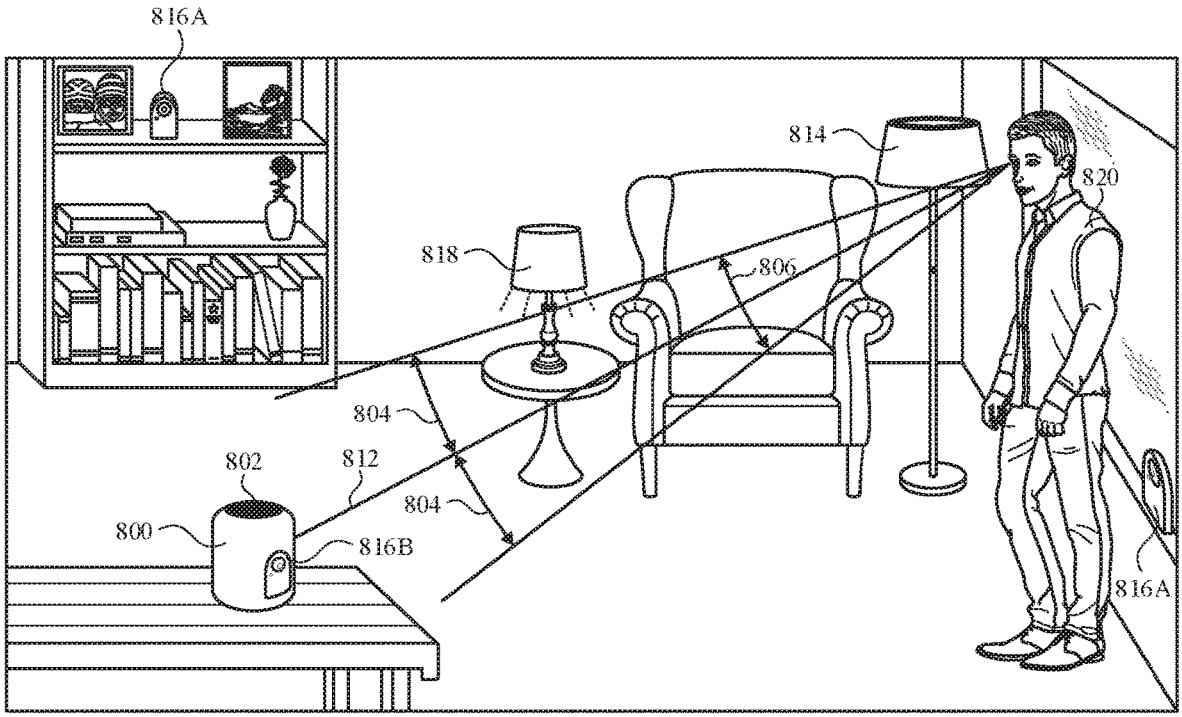

816A
814
820
818
806
804
802
812
804
800
816B
816A

FIG. 8F

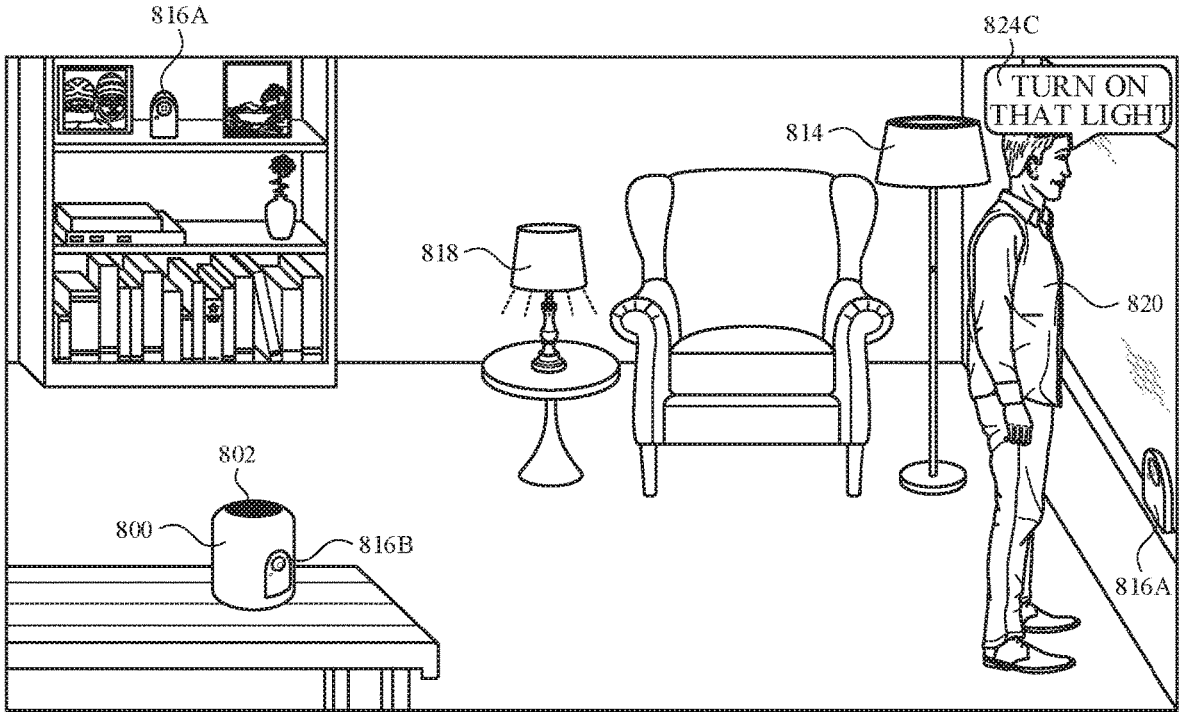

816A
824C
TURN ON THAT LIGHT
814
818
820
802
800
816B
816A

FIG. 8G

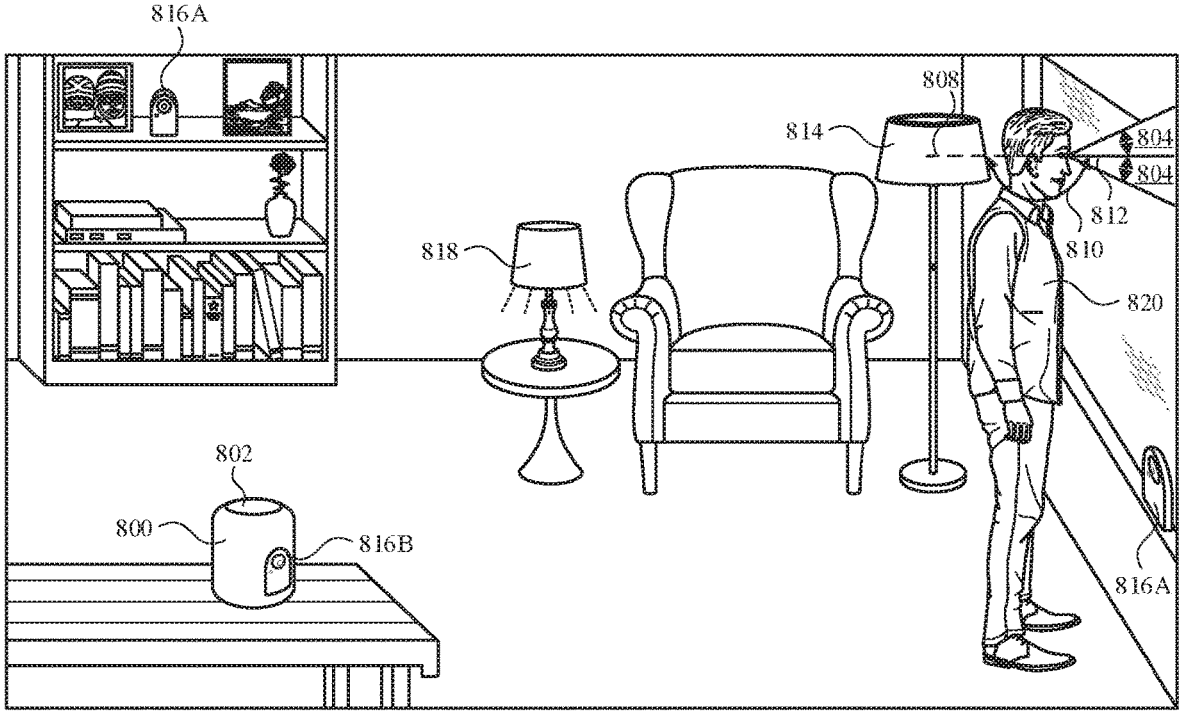

816A
808
814
804
804
812
810
818
820
802
800
816B
816A

FIG. 8H

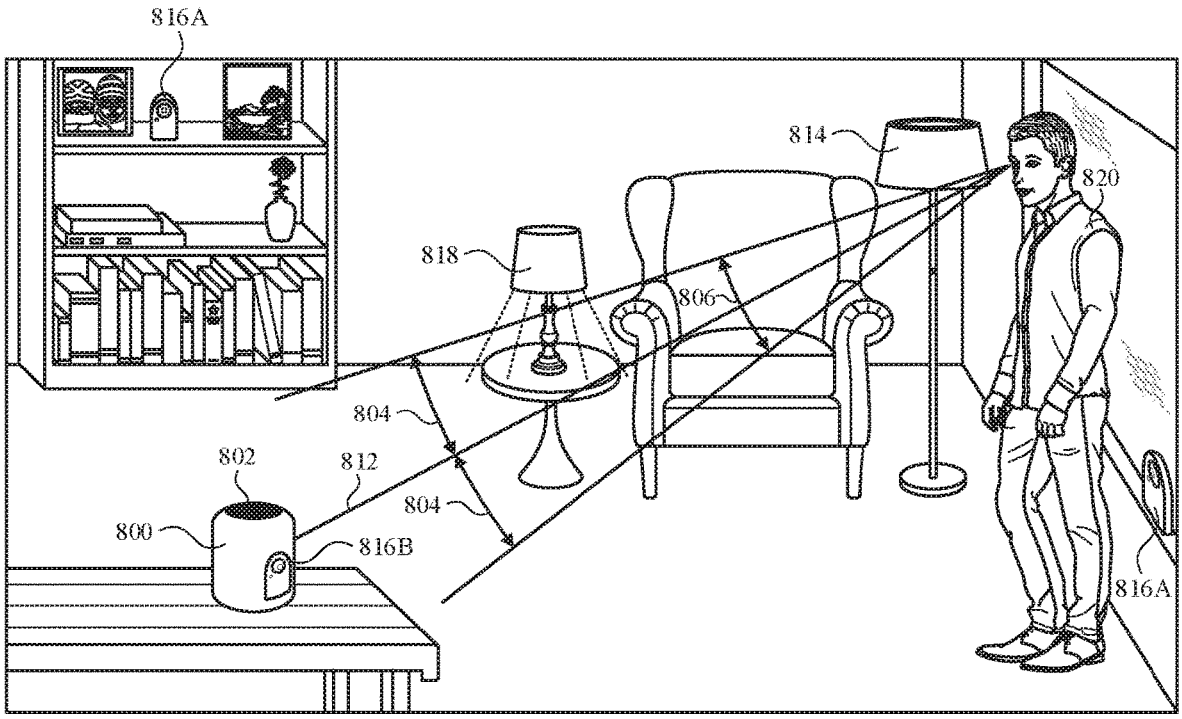

816A
814
820
818
806
804
802
812
804
800
816B
816A

FIG. 8I

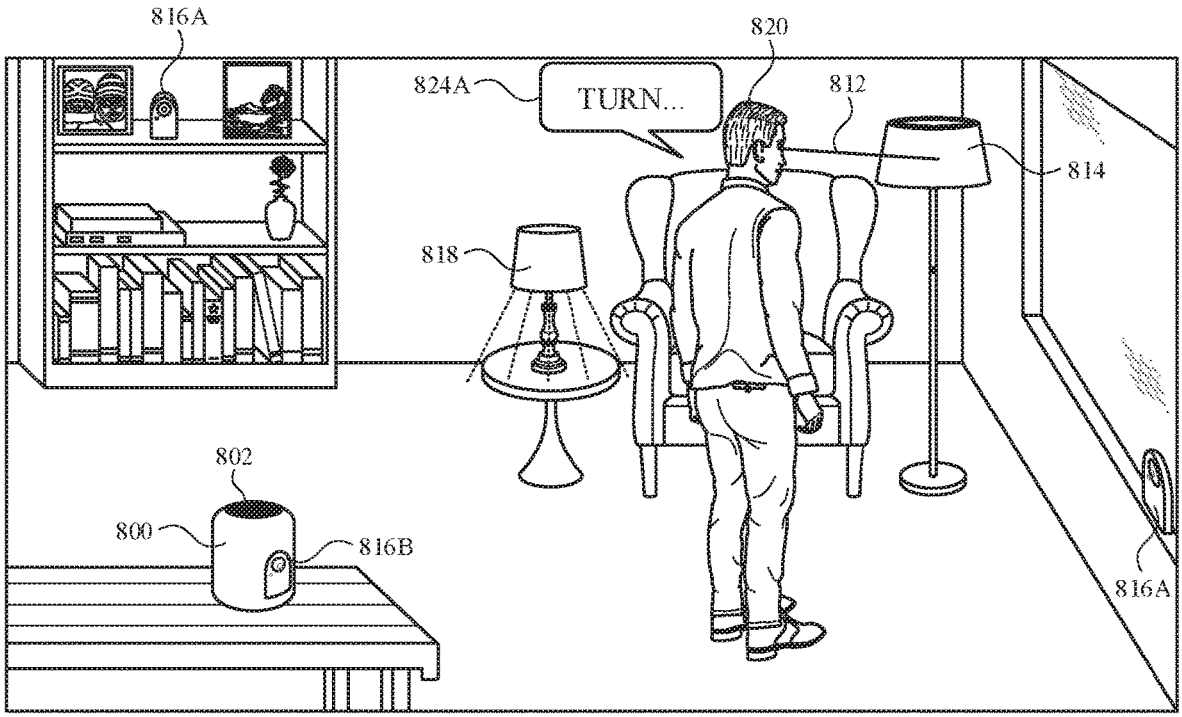

816A
820
824A
TURN...
812
814
818
802
800
816B
816A

FIG. 8J

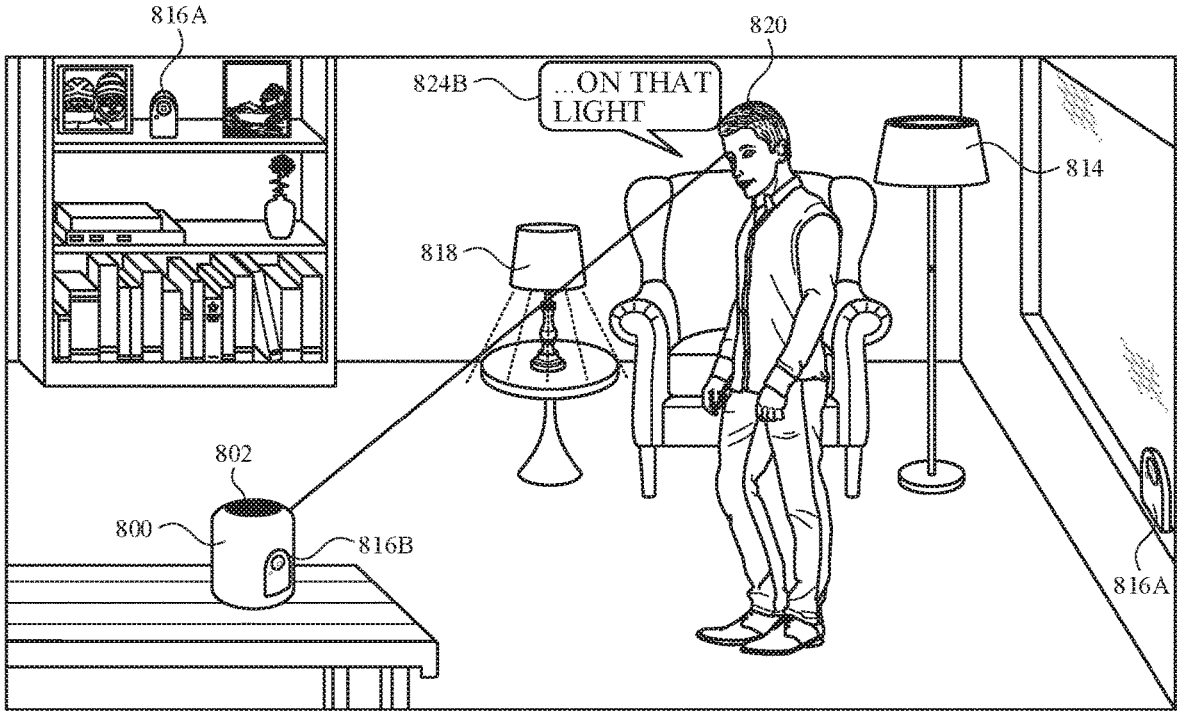

816A
820
824B
...ON THAT LIGHT
814
818
802
800
816B
816A

FIG. 8K

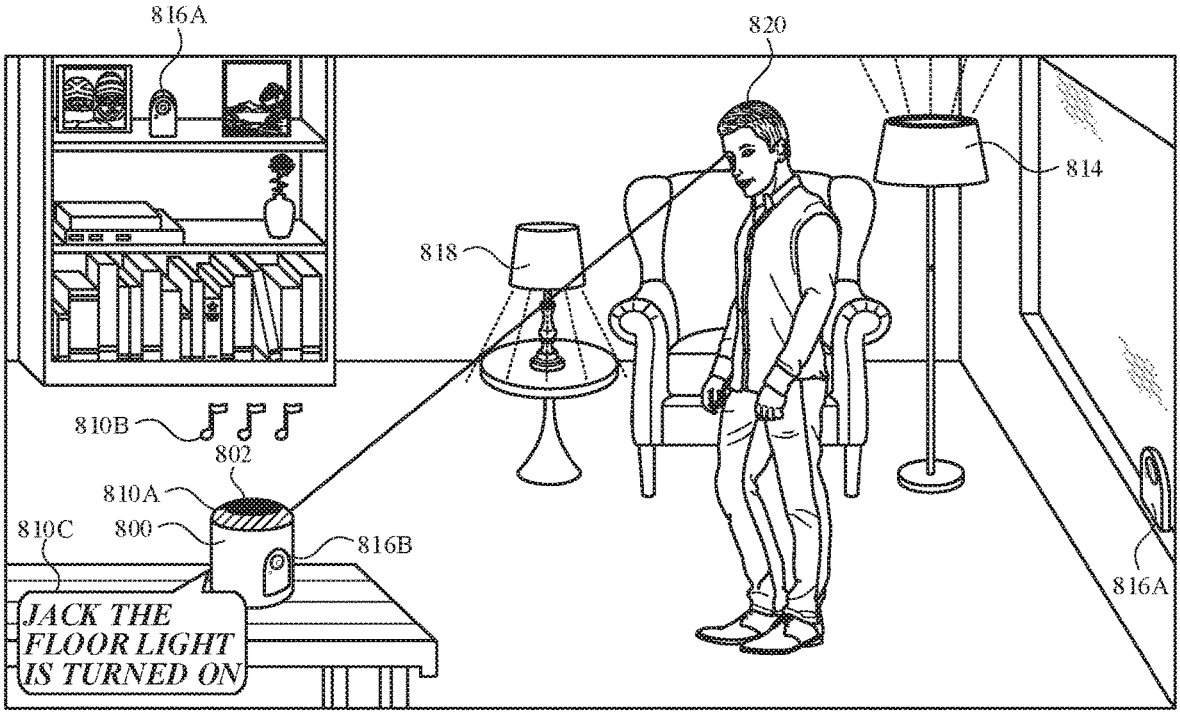

816A
820
814
818
810B
802
810A
810C
800
816B
JACK THE FLOOR LIGHT IS TURNED ON
816A

FIG. 8L

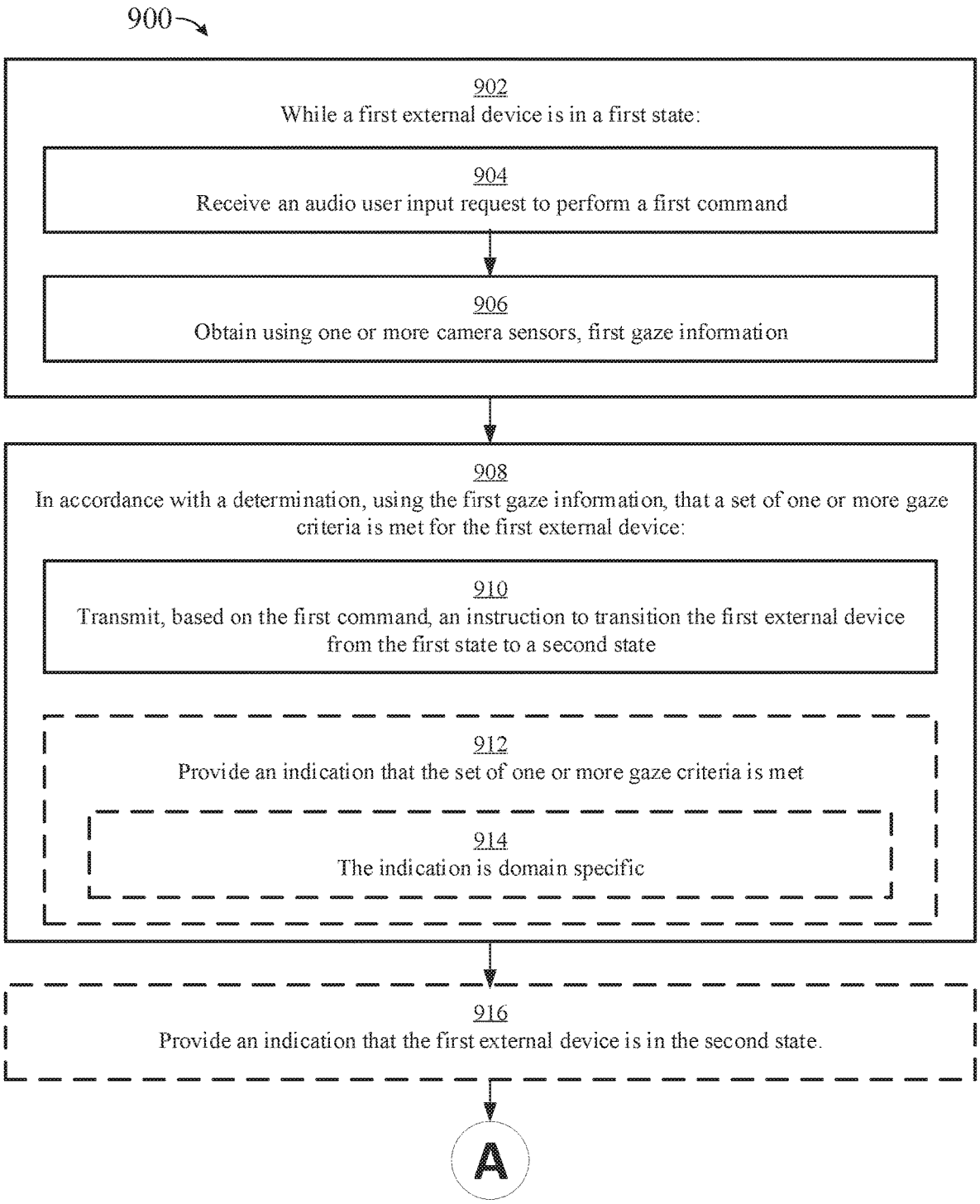

900
902
While a first external device is in a first state:
904
Receive an audio user input request to perform a first command
906
Obtain using one or more camera sensors, first gaze information
908
In accordance with a determination, using the first gaze information, that a set of one or more gaze criteria is met for the first external device:
910
Transmit, based on the first command, an instruction to transition the first external device from the first state to a second state
912
Provide an indication that the set of one or more gaze criteria is met
914
The indication is domain specific
916
Provide an indication that the first external device is in the second state.
A

FIG. 9A

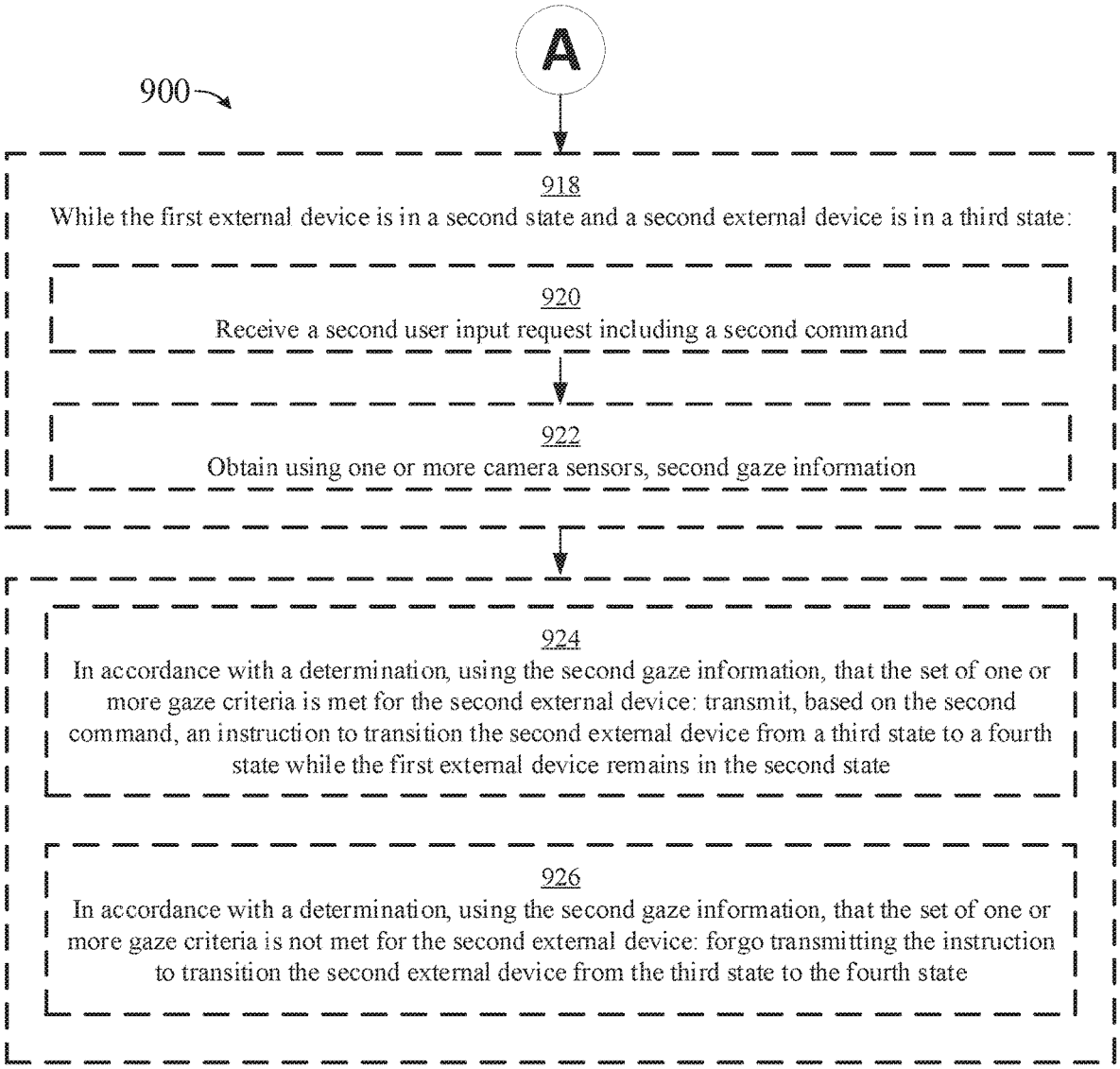

A
900
918
While the first external device is in a second state and a second external device is in a third state:
920
Receive a second user input request including a second command
922
Obtain using one or more camera sensors, second gaze information
924
In accordance with a determination, using the second gaze information, that the set of one or more gaze criteria is met for the second external device: transmit, based on the second command, an instruction to transition the second external device from a third state to a fourth state while the first external device remains in the second state
926
In accordance with a determination, using the second gaze information, that the set of one or more gaze criteria is not met for the second external device: forgo transmitting the instruction to transition the second external device from the third state to the fourth state

FIG. 9B

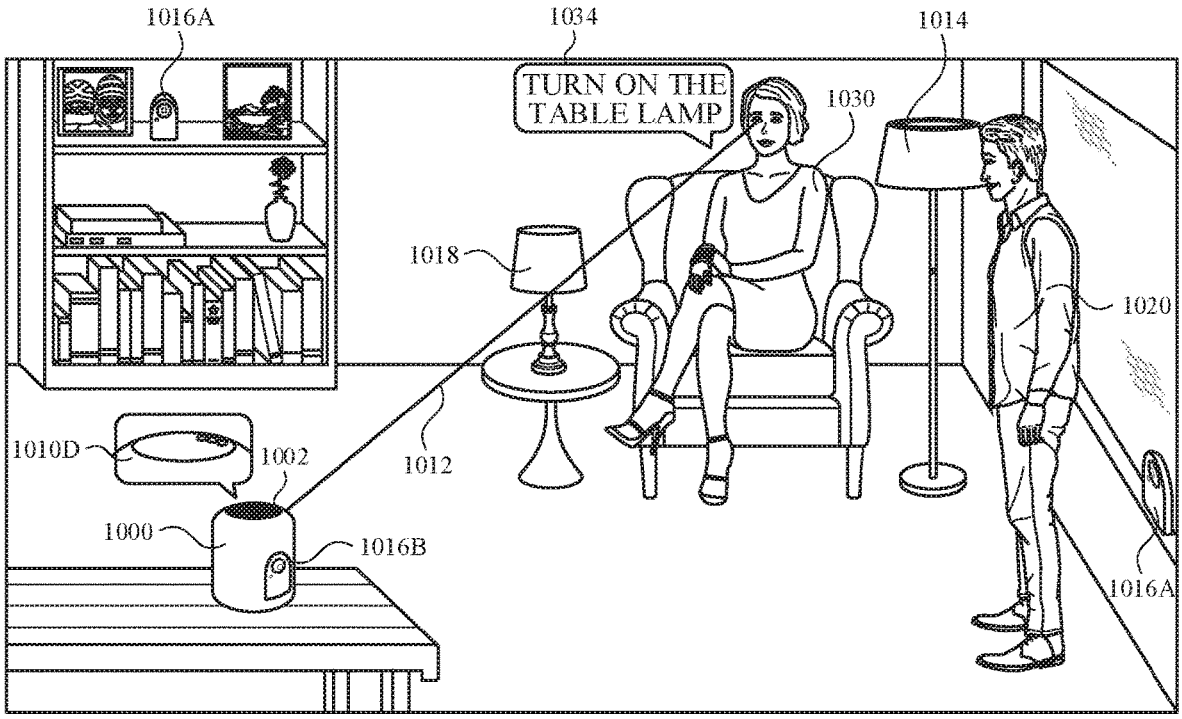


FIG. 10A

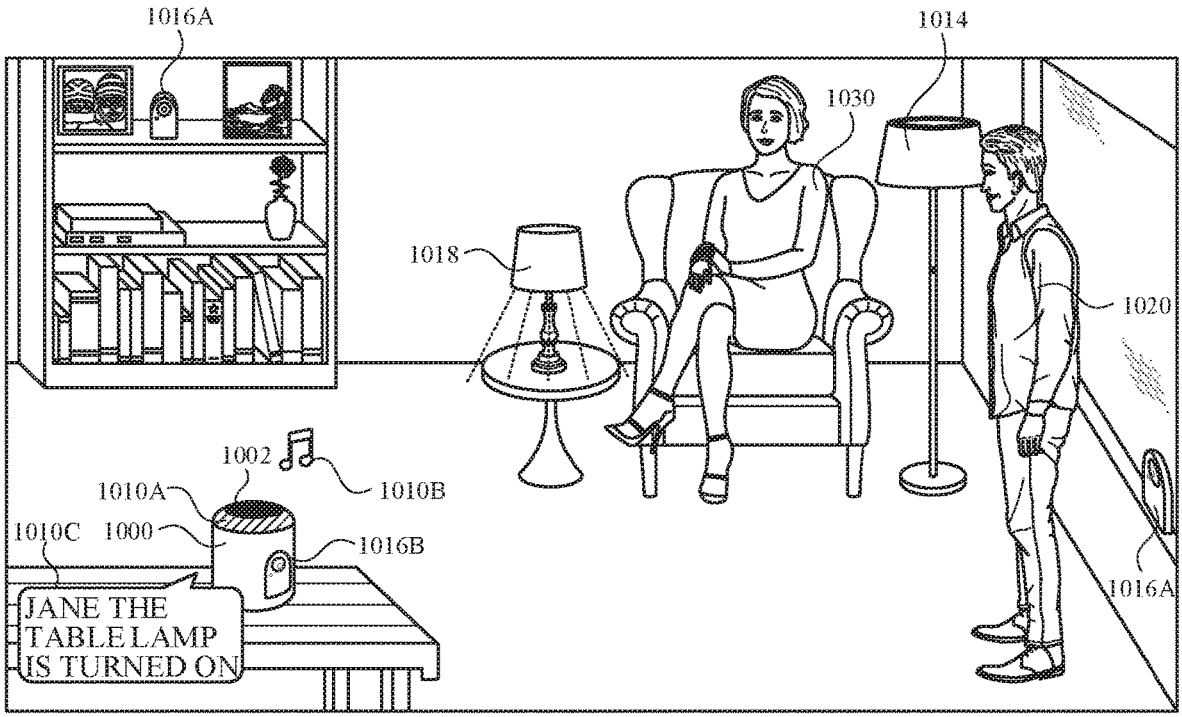


FIG. 10B

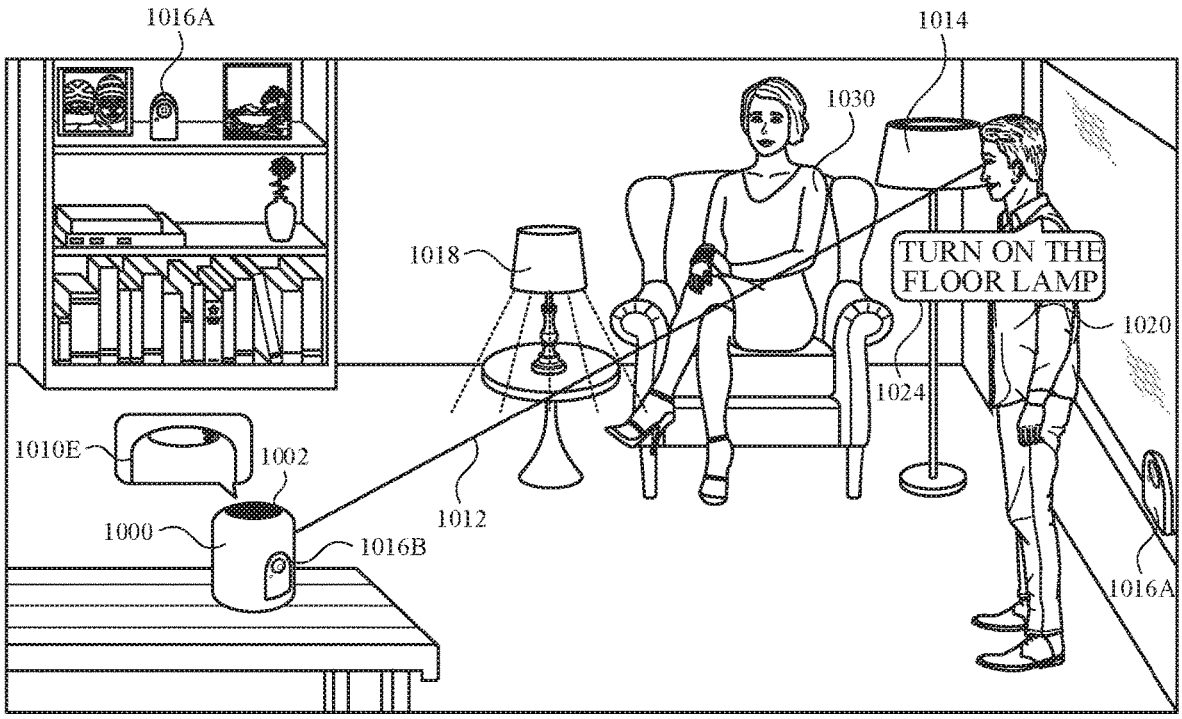


FIG. 10C

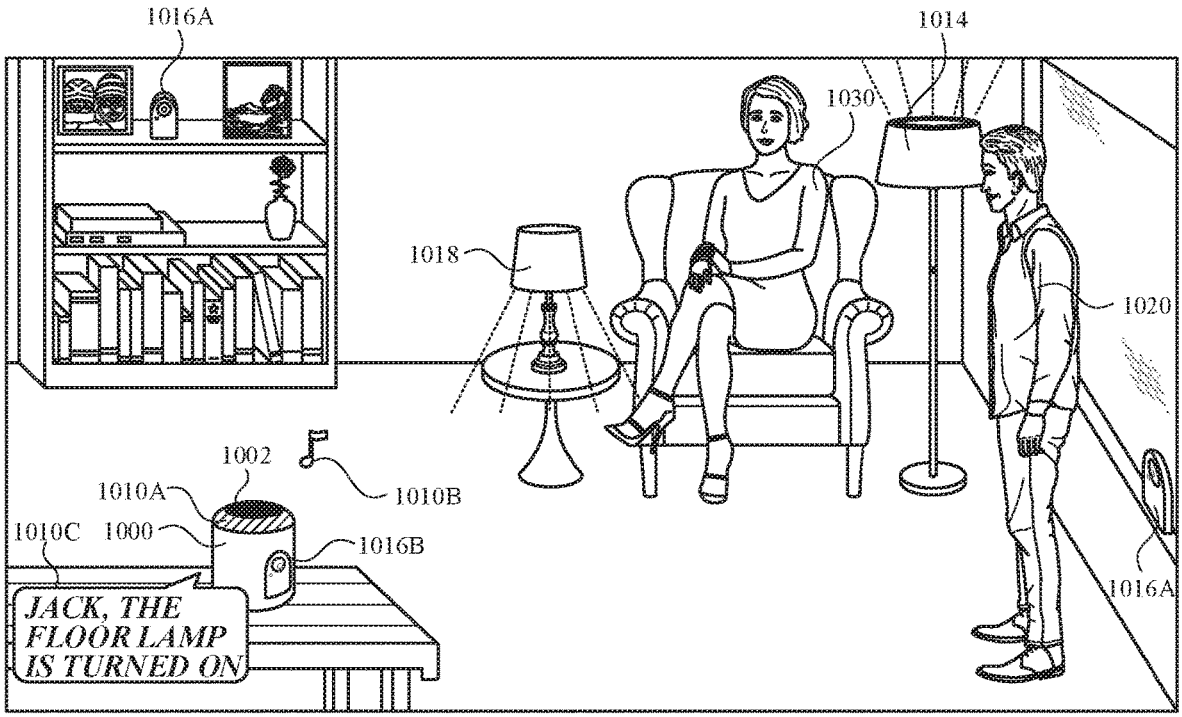


FIG. 10D

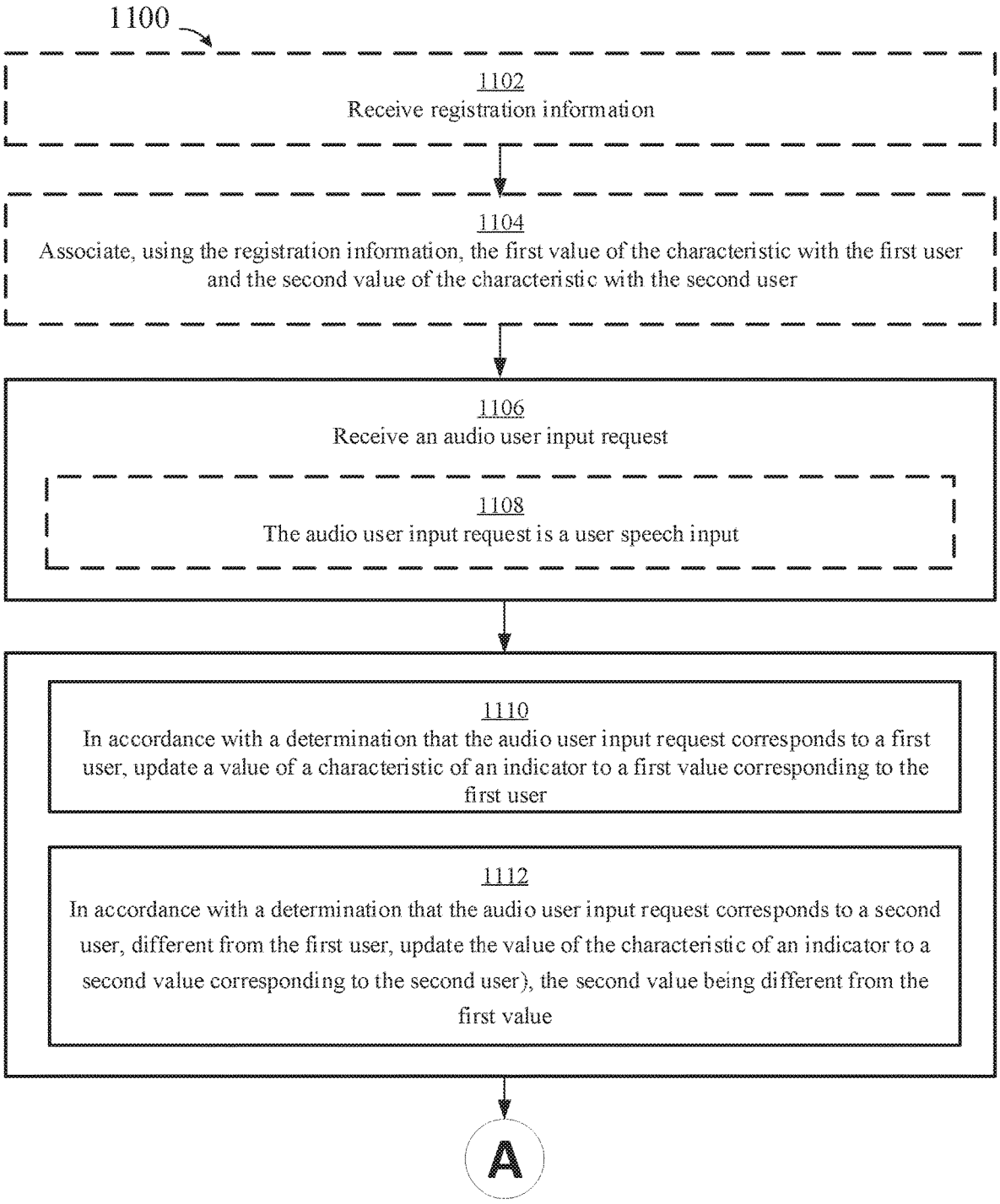

1100
1102
Receive registration information
1104
Associate, using the registration information, the first value of the characteristic with the first user and the second value of the characteristic with the second user
1106
Receive an audio user input request
1108
The audio user input request is a user speech input
1110
In accordance with a determination that the audio user input request corresponds to a first user, update a value of a characteristic of an indicator to a first value corresponding to the first user
1112
In accordance with a determination that the audio user input request corresponds to a second user, different from the first user, update the value of the characteristic of an indicator to a second value corresponding to the second user), the second value being different from the first value
A

FIG. 11A

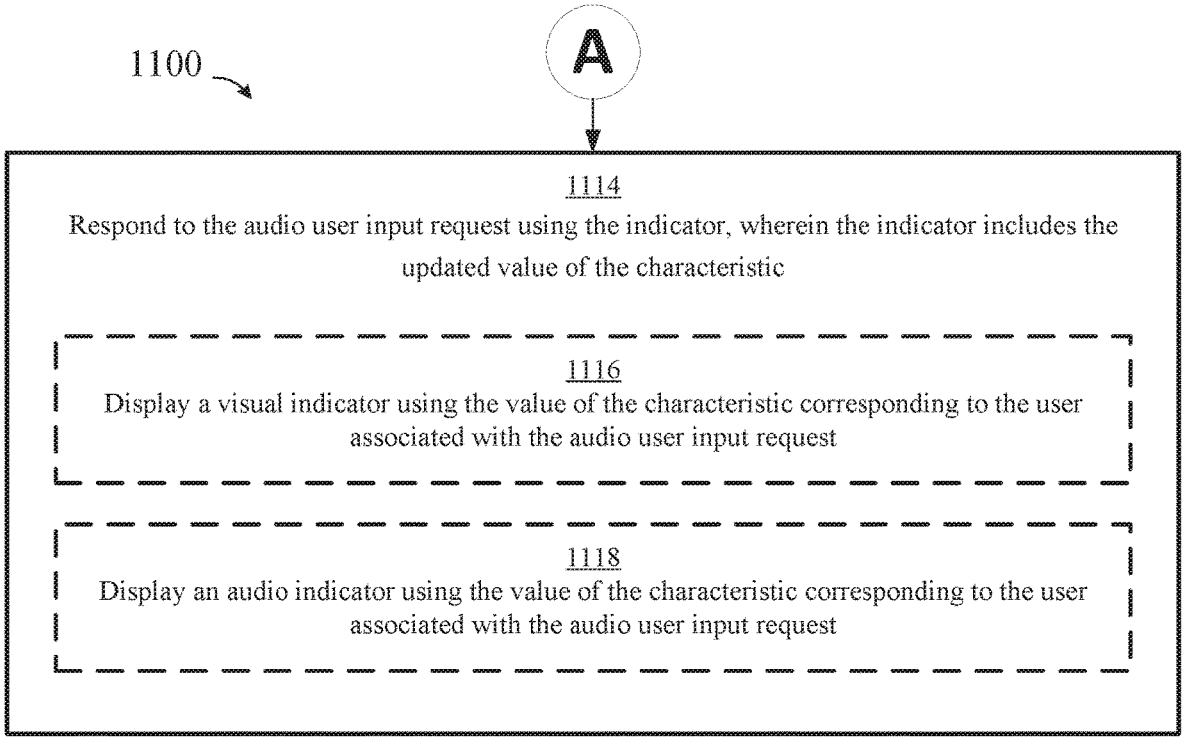

1100
A
1114
Respond to the audio user input request using the indicator, wherein the indicator includes the updated value of the characteristic
1116
Display a visual indicator using the value of the characteristic corresponding to the user associated with the audio user input request
1118
Display an audio indicator using the value of the characteristic corresponding to the user associated with the audio user input request

FIG. 11B

DEVICE CONTROL USING GAZE INFORMATION

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims priority to U.S. Provisional App Ser. No. 62/739,087, entitled "DEVICE CONTROL USING GAZE INFORMATION", filed on Sep. 28, 2018, the content of which is hereby incorporated by reference in its entirety.

FIELD

[0002] The present disclosure relates generally to computer user interfaces, and more specifically to techniques for controlling electronic devices using gaze information.

BACKGROUND

[0003] Users frequently provide inputs, such as key presses and voice inputs, to control electronic devices. For example, users activate a device's button or speak a trigger phrase to start an application on the device. Such inputs frequently require the user to be within arm's reach or within microphone range.

[0004] Intelligent automated assistants (or digital assistants) can provide a beneficial interface between human users and electronic devices. Such assistants can allow users to interact with devices or systems using natural language in spoken and/or text forms. For example, a user can provide a speech input containing a user request to a digital assistant operating on an electronic device. The digital assistant can interpret the user's intent from the speech input, operationalize the user's intent into a task, and perform the task. In some systems, performing tasks in this manner may be constrained in the manner by which a task is identified. In some cases, however, a user may be limited to a particular set of commands such that the user cannot readily instruct a digital assistant to perform a task using natural-language speech inputs. Further, in many instances digital assistants fail to adapt based on previous user behavior and in turn lack a desirable optimization of user experience.

BRIEF SUMMARY

[0005] Some techniques for controlling electronic devices, however, are generally cumbersome and inefficient. For example, some existing techniques use a complex and time-consuming user interface, which may include multiple key presses or keystrokes. For another example, some existing techniques require the user to be within arm's distance to activate a button of the device. Existing techniques require more time than necessary, wasting user time and device energy. This latter consideration is particularly important in battery-operated devices.

[0006] Accordingly, the present technique provides electronic devices with faster, more efficient methods and interfaces for controlling electronic devices. Such methods and interfaces optionally complement or replace other methods for controlling electronic devices. Such methods and interfaces reduce the cognitive burden on a user and produce a more efficient human-machine interface. For battery-operated computing devices, such methods and interfaces conserve power and increase the time between battery charges. Such techniques also allow users to more efficiently interact with electronic devices in environments where the user is not within reaching distance of the electronic device and/or the user is in a noisy environment (e.g., including noise based on audio being produced by the electronic device).

[0007] In accordance with some embodiments, a method is provided. The method is performed at an electronic device. The method comprises: while a digital assistant of the electronic device is not activated: obtaining, using one or more camera sensors, first gaze information; and in accordance with a determination that the first gaze information satisfies a set of one or more activation criteria: activating the digital assistant of the electronic device; and providing an indication that the set of one or more activation criteria has been satisfied.

[0008] In accordance with some embodiments, a non-transitory computer-readable storage medium is provided. The medium stores one or more programs configured to be executed by one or more processors of an electronic device. The one or more programs including instructions for: while a digital assistant of the electronic device is not activated: obtaining, using one or more camera sensors, first gaze information; and in accordance with a determination that the first gaze information satisfies a set of one or more activation criteria: activating the digital assistant of the electronic device; and providing an indication that the set of one or more activation criteria has been satisfied.

[0009] In accordance with some embodiments, a transitory computer-readable storage medium is provided. The medium stores one or more programs configured to be executed by one or more processors of an electronic device. The one or more programs including instructions for: while a digital assistant of the electronic device is not activated: obtaining, using one or more camera sensors, first gaze information; and in accordance with a determination that the first gaze information satisfies a set of one or more activation criteria: activating the digital assistant of the electronic device; and providing an indication that the set of one or more activation criteria has been satisfied.

[0010] In accordance with some embodiments, an electronic device is provided. The electronic device comprises: one or more processors; and memory storing one or more programs configured to be executed by the one or more processors, the one or more programs including instructions for: while a digital assistant of the electronic device is not activated: obtaining, using one or more camera sensors, first gaze information; and in accordance with a determination that the first gaze information satisfies a set of one or more activation criteria: activating the digital assistant of the electronic device; and providing an indication that the set of one or more activation criteria has been satisfied.

[0011] In accordance with some embodiments, an electronic device is provided. The electronic device comprises: means, while a digital assistant of the electronic device is not activated, for: obtaining, using one or more camera sensors, first gaze information; and means, in accordance with a determination that the first gaze information satisfies a set of one or more activation criteria, for: activating the digital assistant of the electronic device; and providing an indication that the set of one or more activation criteria has been satisfied.

[0012] In accordance with some embodiments, a method is provided. The method is performed at an electronic device. The method comprises: while a first external device is in a first state: receiving an audio user input request to perform a first command; and obtaining, using one or more

camera sensors, first gaze information; and in accordance with a determination, using the first gaze information, that a set of one or more gaze criteria is met for the first external device: transmitting, based on the first command, an instruction to transition the first external device from the first state to a second state.

[0013] In accordance with some embodiments, a non-transitory computer-readable storage medium is provided. The medium stores one or more programs configured to be executed by one or more processors of an electronic device. The one or more programs including instructions for: while a first external device is in a first state: receiving an audio user input request to perform a first command; and obtaining, using one or more camera sensors, first gaze information; and in accordance with a determination, using the first gaze information, that a set of one or more gaze criteria is met for the first external device: transmitting, based on the first command, an instruction to transition the first external device from the first state to a second state.

[0014] In accordance with some embodiments, a transitory computer-readable storage medium is provided. The medium stores one or more programs configured to be executed by one or more processors of an electronic device. The one or more programs including instructions for: while a first external device is in a first state: receiving an audio user input request to perform a first command; and obtaining, using one or more camera sensors, first gaze information; and in accordance with a determination, using the first gaze information, that a set of one or more gaze criteria is met for the first external device: transmitting, based on the first command, an instruction to transition the first external device from the first state to a second state.

[0015] In accordance with some embodiments, an electronic device is provided. The electronic device comprises: one or more processors; and memory storing one or more programs configured to be executed by the one or more processors, the one or more programs including instructions for: while a first external device is in a first state: receiving an audio user input request to perform a first command; and obtaining, using one or more camera sensors, first gaze information; and in accordance with a determination, using the first gaze information, that a set of one or more gaze criteria is met for the first external device: transmitting, based on the first command, an instruction to transition the first external device from the first state to a second state.

[0016] In accordance with some embodiments, an electronic device is provided. The electronic device comprises: means, while a first external device is in a first state, for: receiving an audio user input request to perform a first command; and obtaining, using one or more camera sensors, first gaze information; and means, in accordance with a determination, using the first gaze information, that a set of one or more gaze criteria is met for the first external device, for: transmitting, based on the first command, an instruction to transition the first external device from the first state to a second state.

[0017] In accordance with some embodiments, a method is provided. The method is performed at an electronic device. The method comprises: receiving an audio user input request; in accordance with a determination that the audio user input request corresponds to a first user, updating a value of a characteristic of an indicator to a first value corresponding to the first user; in accordance with a determination that the audio user input request corresponds to a second user, different from the first user, updating the value of the characteristic of the indicator to a second value corresponding to the second user, the second value being different from the first value; and responding to the audio user input request using the indicator, wherein the indicator includes the updated value of the characteristic.

[0018] In accordance with some embodiments, a non-transitory computer-readable storage medium is provided. The medium stores one or more programs configured to be executed by one or more processors of an electronic device. The one or more programs including instructions for: receiving an audio user input request; in accordance with a determination that the audio user input request corresponds to a first user, updating a value of a characteristic of an indicator to a first value corresponding to the first user; in accordance with a determination that the audio user input request corresponds to a second user, different from the first user, updating the value of the characteristic of the indicator to a second value corresponding to the second user, the second value being different from the first value; and responding to the audio user input request using the indicator, wherein the indicator includes the updated value of the characteristic.

[0019] In accordance with some embodiments, a transitory computer-readable storage medium is provided. The medium stores one or more programs configured to be executed by one or more processors of an electronic device. The one or more programs including instructions for: receiving an audio user input request; in accordance with a determination that the audio user input request corresponds to a first user, updating a value of a characteristic of an indicator to a first value corresponding to the first user; in accordance with a determination that the audio user input request corresponds to a second user, different from the first user, updating the value of the characteristic of the indicator to a second value corresponding to the second user, the second value being different from the first value; and responding to the audio user input request using the indicator, wherein the indicator includes the updated value of the characteristic.

[0020] In accordance with some embodiments, an electronic device is provided. The electronic device comprises: one or more processors; and memory storing one or more programs configured to be executed by the one or more processors, the one or more programs including instructions for: receiving an audio user input request; in accordance with a determination that the audio user input request corresponds to a first user, updating a value of a characteristic of an indicator to a first value corresponding to the first user; in accordance with a determination that the audio user input request corresponds to a second user, different from the first user, updating the value of the characteristic of the indicator to a second value corresponding to the second user, the second value being different from the first value; and responding to the audio user input request using the indicator, wherein the indicator includes the updated value of the characteristic.

[0021] In accordance with some embodiments, an electronic device is provided. The electronic device comprises: means for receiving an audio user input request; means, in accordance with a determination that the audio user input request corresponds to a first user, for updating a value of a characteristic of an indicator to a first value corresponding to the first user; means, in accordance with a determination that

the audio user input request corresponds to a second user, different from the first user, for updating the value of the characteristic of the indicator to a second value corresponding to the second user, the second value being different from the first value; and means for responding to the audio user input request using the indicator, wherein the indicator includes the updated value of the characteristic.

[0022] Executable instructions for performing these functions are, optionally, included in a non-transitory computer-readable storage medium or other computer program product configured for execution by one or more processors. Executable instructions for performing these functions are, optionally, included in a transitory computer-readable storage medium or other computer program product configured for execution by one or more processors.

[0023] Thus, devices are provided with faster, more efficient methods and interfaces for controlling electronic devices, thereby increasing the effectiveness, efficiency, and user satisfaction with such devices. Such methods and interfaces may complement or replace other methods for controlling electronic devices.

DESCRIPTION OF THE FIGURES

[0024] For a better understanding of the various described embodiments, reference should be made to the Description of Embodiments below, in conjunction with the following drawings in which like reference numerals refer to corresponding parts throughout the figures.

[0025] FIG. **1**A is a block diagram illustrating a portable multifunction device with a touch-sensitive display in accordance with some embodiments.

[0026] FIG. **1**B is a block diagram illustrating exemplary components for event handling in accordance with some embodiments.

[0027] FIG. **1**C is a block diagram illustrating a system and environment for implementing a digital assistant, in accordance with some embodiments.

[0028] FIG. **2** illustrates a portable multifunction device having a touch screen in accordance with some embodiments.

[0029] FIG. **3** is a block diagram of an exemplary multifunction device with a display and a touch-sensitive surface in accordance with some embodiments.

[0030] FIG. **4**A illustrates an exemplary user interface for a menu of applications on a portable multifunction device in accordance with some embodiments.

[0031] FIG. **4**B illustrates an exemplary user interface for a multifunction device with a touch-sensitive surface that is separate from the display in accordance with some embodiments.

[0032] FIG. **5**A illustrates a personal electronic device in accordance with some embodiments.

[0033] FIG. **5**B is a block diagram illustrating a personal electronic device in accordance with some embodiments.

[0034] FIGS. **6**A-**6**O illustrate exemplary techniques for activating a digital assistant using gaze information.

[0035] FIGS. **7**A-**7**B are flow diagrams illustrating a method for activating a digital assistant using gaze information.

[0036] FIGS. **8**A-**8**L illustrate exemplary techniques for providing context to commands using gaze information.

[0037] FIGS. **9**A-**9**B are flow diagrams illustrating a method for providing context to commands using gaze information.

[0038] FIGS. **10**A-**10**D illustrate exemplary techniques for providing indications that distinguish between different speakers.

[0039] FIGS. **11**A-**11**B are flow diagrams illustrating a method for providing indications that distinguish between different speakers.

DESCRIPTION OF EMBODIMENTS

[0040] The following description sets forth exemplary methods, parameters, and the like. It should be recognized, however, that such description is not intended as a limitation on the scope of the present disclosure but is instead provided as a description of exemplary embodiments.

[0041] There is a need for electronic devices that provide efficient methods and interfaces for controlling electronic devices. For example, eliminating or reducing the need for users to provide tactile or verbal input to activate a digital assistant enables users to more effectively control electronic devices. For another example, identifying an external device through the user's gaze eliminates or reduces the need to provide complex and time-consuming user input to identify such a device. Such techniques can reduce the cognitive burden on a user who control electronic devices, thereby enhancing productivity. Further, such techniques can reduce processor and battery power otherwise wasted on redundant user inputs.

[0042] Below, FIGS. **1**A-**1**C, **2**, **3**, **4**A-**4**B, and **5**A-**5**B provide a description of exemplary devices for performing the techniques for controlling electronic devices. FIGS. **6**A-**6**O illustrate exemplary user interfaces for activating a digital assistant using gaze information. FIGS. **7**A-**7**B are a flow diagram illustrating methods of activating a digital assistant using gaze information in accordance with some embodiments. The user interfaces in FIGS. **6**A-**6**O are used to illustrate the processes described below, including the processes in FIGS. **7**A-**7**B. FIGS. **8**A-**8**L illustrate exemplary user interfaces for providing context to commands using gaze information. FIGS. **9**A-**9**B are a flow diagram illustrating methods of providing context to commands using gaze information in accordance with some embodiments. The user interfaces in FIGS. **8**A-**8**L are used to illustrate the processes described below, including the processes in FIGS. **9**A-**9**B. FIGS. **10**A-**10**D illustrate exemplary user interfaces for providing indications that distinguish between different speakers. FIGS. **11**A-**11**B are a flow diagram illustrating methods of providing indications that distinguish between different speakers in accordance with some embodiments. The user interfaces in FIGS. **10**A-**10**D are used to illustrate the processes described below, including the processes in FIGS. **11**A-**11**B.

[0043] Although the following description uses terms "first," "second," etc. to describe various elements, these elements should not be limited by the terms. These terms are only used to distinguish one element from another. For example, a first touch could be termed a second touch, and, similarly, a second touch could be termed a first touch, without departing from the scope of the various described embodiments. The first touch and the second touch are both touches, but they are not the same touch.

[0044] The terminology used in the description of the various described embodiments herein is for the purpose of describing particular embodiments only and is not intended to be limiting. As used in the description of the various described embodiments and the appended claims, the sin-

gular forms "a," "an," and "the" are intended to include the plural forms as well, unless the context clearly indicates otherwise. It will also be understood that the term "and/or" as used herein refers to and encompasses any and all possible combinations of one or more of the associated listed items. It will be further understood that the terms "includes," "including," "comprises," and/or "comprising," when used in this specification, specify the presence of stated features, integers, steps, operations, elements, and/or components, but do not preclude the presence or addition of one or more other features, integers, steps, operations, elements, components, and/or groups thereof.

[0045] The term "if" is, optionally, construed to mean "when" or "upon" or "in response to determining" or "in response to detecting," depending on the context. Similarly, the phrase "if it is determined" or "if [a stated condition or event] is detected" is, optionally, construed to mean "upon determining" or "in response to determining" or "upon detecting [the stated condition or event]" or "in response to detecting [the stated condition or event]," depending on the context.

[0046] Embodiments of electronic devices, user interfaces for such devices, and associated processes for using such devices are described. In some embodiments, the device is a portable communications device, such as a mobile telephone, that also contains other functions, such as PDA and/or music player functions. Exemplary embodiments of portable multifunction devices include, without limitation, the iPhone®, iPod Touch®, and iPad® devices from Apple Inc. of Cupertino, Calif. Other portable electronic devices, such as laptops or tablet computers with touch-sensitive surfaces (e.g., touch screen displays and/or touchpads), are, optionally, used. It should also be understood that, in some embodiments, the device is not a portable communications device, but is a desktop computer with a touch-sensitive surface (e.g., a touch screen display and/or a touchpad).

[0047] In the discussion that follows, an electronic device that includes a display and a touch-sensitive surface is described. It should be understood, however, that the electronic device optionally includes one or more other physical user-interface devices, such as a physical keyboard, a mouse, and/or a joystick.

[0048] The device typically supports a variety of applications, such as one or more of the following: a drawing application, a presentation application, a word processing application, a website creation application, a disk authoring application, a spreadsheet application, a gaming application, a telephone application, a video conferencing application, an e-mail application, an instant messaging application, a workout support application, a photo management application, a digital camera application, a digital video camera application, a web browsing application, a digital music player application, and/or a digital video player application.

[0049] The various applications that are executed on the device optionally use at least one common physical user-interface device, such as the touch-sensitive surface. One or more functions of the touch-sensitive surface as well as corresponding information displayed on the device are, optionally, adjusted and/or varied from one application to the next and/or within a respective application. In this way, a common physical architecture (such as the touch-sensitive surface) of the device optionally supports the variety of applications with user interfaces that are intuitive and transparent to the user.

[0050] Attention is now directed toward embodiments of portable devices with touch-sensitive displays. FIG. **1**A is a block diagram illustrating portable multifunction device **100** with touch-sensitive display system **112** in accordance with some embodiments. Touch-sensitive display **112** is sometimes called a "touch screen" for convenience and is sometimes known as or called a "touch-sensitive display system." Device **100** includes memory **102** (which optionally includes one or more computer-readable storage mediums), memory controller **122**, one or more processing units (CPUs) **120**, peripherals interface **118**, RF circuitry **108**, audio circuitry **110**, speaker **111**, microphone **113**, input/output (I/O) subsystem **106**, other input control devices **116**, and external port **124**. Device **100** optionally includes one or more optical sensors **164**. Device **100** optionally includes one or more contact intensity sensors **165** for detecting intensity of contacts on device **100** (e.g., a touch-sensitive surface such as touch-sensitive display system **112** of device **100**). Device **100** optionally includes one or more tactile output generators **167** for generating tactile outputs on device **100** (e.g., generating tactile outputs on a touch-sensitive surface such as touch-sensitive display system **112** of device **100** or touchpad **355** of device **300**). These components optionally communicate over one or more communication buses or signal lines **103**.

[0051] As used in the specification and claims, the term "intensity" of a contact on a touch-sensitive surface refers to the force or pressure (force per unit area) of a contact (e.g., a finger contact) on the touch-sensitive surface, or to a substitute (proxy) for the force or pressure of a contact on the touch-sensitive surface. The intensity of a contact has a range of values that includes at least four distinct values and more typically includes hundreds of distinct values (e.g., at least 256). Intensity of a contact is, optionally, determined (or measured) using various approaches and various sensors or combinations of sensors. For example, one or more force sensors underneath or adjacent to the touch-sensitive surface are, optionally, used to measure force at various points on the touch-sensitive surface. In some implementations, force measurements from multiple force sensors are combined (e.g., a weighted average) to determine an estimated force of a contact. Similarly, a pressure-sensitive tip of a stylus is, optionally, used to determine a pressure of the stylus on the touch-sensitive surface. Alternatively, the size of the contact area detected on the touch-sensitive surface and/or changes thereto, the capacitance of the touch-sensitive surface proximate to the contact and/or changes thereto, and/or the resistance of the touch-sensitive surface proximate to the contact and/or changes thereto are, optionally, used as a substitute for the force or pressure of the contact on the touch-sensitive surface. In some implementations, the substitute measurements for contact force or pressure are used directly to determine whether an intensity threshold has been exceeded (e.g., the intensity threshold is described in units corresponding to the substitute measurements). In some implementations, the substitute measurements for contact force or pressure are converted to an estimated force or pressure, and the estimated force or pressure is used to determine whether an intensity threshold has been exceeded (e.g., the intensity threshold is a pressure threshold measured in units of pressure). Using the intensity of a contact as an attribute of a user input allows for user access to additional device functionality that may otherwise not be accessible by the user on a reduced-size device with limited

real estate for displaying affordances (e.g., on a touch-sensitive display) and/or receiving user input (e.g., via a touch-sensitive display, a touch-sensitive surface, or a physical/mechanical control such as a knob or a button).

[0052] As used in the specification and claims, the term "tactile output" refers to physical displacement of a device relative to a previous position of the device, physical displacement of a component (e.g., a touch-sensitive surface) of a device relative to another component (e.g., housing) of the device, or displacement of the component relative to a center of mass of the device that will be detected by a user with the user's sense of touch. For example, in situations where the device or the component of the device is in contact with a surface of a user that is sensitive to touch (e.g., a finger, palm, or other part of a user's hand), the tactile output generated by the physical displacement will be interpreted by the user as a tactile sensation corresponding to a perceived change in physical characteristics of the device or the component of the device. For example, movement of a touch-sensitive surface (e.g., a touch-sensitive display or trackpad) is, optionally, interpreted by the user as a "down click" or "up click" of a physical actuator button. In some cases, a user will feel a tactile sensation such as an "down click" or "up click" even when there is no movement of a physical actuator button associated with the touch-sensitive surface that is physically pressed (e.g., displaced) by the user's movements. As another example, movement of the touch-sensitive surface is, optionally, interpreted or sensed by the user as "roughness" of the touch-sensitive surface, even when there is no change in smoothness of the touch-sensitive surface. While such interpretations of touch by a user will be subject to the individualized sensory perceptions of the user, there are many sensory perceptions of touch that are common to a large majority of users. Thus, when a tactile output is described as corresponding to a particular sensory perception of a user (e.g., an "up click," a "down click," "roughness"), unless otherwise stated, the generated tactile output corresponds to physical displacement of the device or a component thereof that will generate the described sensory perception for a typical (or average) user.

[0053] It should be appreciated that device **100** is only one example of a portable multifunction device, and that device **100** optionally has more or fewer components than shown, optionally combines two or more components, or optionally has a different configuration or arrangement of the components. The various components shown in FIG. **1**A are implemented in hardware, software, or a combination of both hardware and software, including one or more signal processing and/or application-specific integrated circuits.

[0054] Memory **102** optionally includes high-speed random access memory and optionally also includes non-volatile memory, such as one or more magnetic disk storage devices, flash memory devices, or other non-volatile solid-state memory devices. Memory controller **122** optionally controls access to memory **102** by other components of device **100**.

[0055] Peripherals interface **118** can be used to couple input and output peripherals of the device to CPU **120** and memory **102**. The one or more processors **120** run or execute various software programs and/or sets of instructions stored in memory **102** to perform various functions for device **100** and to process data. In some embodiments, peripherals interface **118**, CPU **120**, and memory controller **122** are, optionally, implemented on a single chip, such as chip **104**. In some other embodiments, they are, optionally, implemented on separate chips.

[0056] RF (radio frequency) circuitry **108** receives and sends RF signals, also called electromagnetic signals. RF circuitry **108** converts electrical signals to/from electromagnetic signals and communicates with communications networks and other communications devices via the electromagnetic signals. RF circuitry **108** optionally includes well-known circuitry for performing these functions, including but not limited to an antenna system, an RF transceiver, one or more amplifiers, a tuner, one or more oscillators, a digital signal processor, a CODEC chipset, a subscriber identity module (SIM) card, memory, and so forth. RF circuitry **108** optionally communicates with networks, such as the Internet, also referred to as the World Wide Web (WWW), an intranet and/or a wireless network, such as a cellular telephone network, a wireless local area network (LAN) and/or a metropolitan area network (MAN), and other devices by wireless communication. The RF circuitry **108** optionally includes well-known circuitry for detecting near field communication (NFC) fields, such as by a short-range communication radio. The wireless communication optionally uses any of a plurality of communications standards, protocols, and technologies, including but not limited to Global System for Mobile Communications (GSM), Enhanced Data GSM Environment (EDGE), high-speed downlink packet access (HSDPA), high-speed uplink packet access (HSUPA), Evolution, Data-Only (EV-DO), HSPA, HSPA+, Dual-Cell HSPA (DC-HSPDA), long term evolution (LTE), near field communication (NFC), wideband code division multiple access (W-CDMA), code division multiple access (CDMA), time division multiple access (TDMA), Bluetooth, Bluetooth Low Energy (BTLE), Wireless Fidelity (Wi-Fi) (e.g., IEEE 802.11a, IEEE 802.11b, IEEE 802.11g, IEEE 802.11n, and/or IEEE 802.11ac), voice over Internet Protocol (VoIP), Wi-MAX, a protocol for e-mail (e.g., Internet message access protocol (IMAP) and/or post office protocol (POP)), instant messaging (e.g., extensible messaging and presence protocol (XMPP), Session Initiation Protocol for Instant Messaging and Presence Leveraging Extensions (SIMPLE), Instant Messaging and Presence Service (IMPS)), and/or Short Message Service (SMS), or any other suitable communication protocol, including communication protocols not yet developed as of the filing date of this document.

[0057] Audio circuitry **110**, speaker **111**, and microphone **113** provide an audio interface between a user and device **100**. Audio circuitry **110** receives audio data from peripherals interface **118**, converts the audio data to an electrical signal, and transmits the electrical signal to speaker **111**. Speaker **111** converts the electrical signal to human-audible sound waves. Audio circuitry **110** also receives electrical signals converted by microphone **113** from sound waves. Audio circuitry **110** converts the electrical signal to audio data and transmits the audio data to peripherals interface **118** for processing. Audio data is, optionally, retrieved from and/or transmitted to memory **102** and/or RF circuitry **108** by peripherals interface **118**. In some embodiments, audio circuitry **110** also includes a headset jack (e.g., **212**, FIG. **2**). The headset jack provides an interface between audio circuitry **110** and removable audio input/output peripherals, such as output-only headphones or a headset with both output (e.g., a headphone for one or both ears) and input (e.g., a microphone).

[0058] I/O subsystem **106** couples input/output peripherals on device **100**, such as touch screen **112** and other input control devices **116**, to peripherals interface **118**. I/O subsystem **106** optionally includes display controller **156**, optical sensor controller **158**, depth camera controller **169**, intensity sensor controller **159**, haptic feedback controller **161**, and one or more input controllers **160** for other input or control devices. The one or more input controllers **160** receive/send electrical signals from/to other input control devices **116**. The other input control devices **116** optionally include physical buttons (e.g., push buttons, rocker buttons, etc.), dials, slider switches, joysticks, click wheels, and so forth. In some alternate embodiments, input controller(s) **160** are, optionally, coupled to any (or none) of the following: a keyboard, an infrared port, a USB port, and a pointer device such as a mouse. The one or more buttons (e.g., **208**, FIG. **2**) optionally include an up/down button for volume control of speaker **111** and/or microphone **113**. The one or more buttons optionally include a push button (e.g., **206**, FIG. **2**).

[0059] A quick press of the push button optionally disengages a lock of touch screen **112** or optionally begins a process that uses gestures on the touch screen to unlock the device, as described in U.S. patent application Ser. No. 11/322,549, "Unlocking a Device by Performing Gestures on an Unlock Image," filed Dec. 23, 2005, U.S. Pat. No. 7,657,849, which is hereby incorporated by reference in its entirety. A longer press of the push button (e.g., **206**) optionally turns power to device **100** on or off. The functionality of one or more of the buttons are, optionally, user-customizable. Touch screen **112** is used to implement virtual or soft buttons and one or more soft keyboards.

[0060] Touch-sensitive display **112** provides an input interface and an output interface between the device and a user. Display controller **156** receives and/or sends electrical signals from/to touch screen **112**. Touch screen **112** displays visual output to the user. The visual output optionally includes graphics, text, icons, video, and any combination thereof (collectively termed "graphics"). In some embodiments, some or all of the visual output optionally corresponds to user-interface objects.

[0061] Touch screen **112** has a touch-sensitive surface, sensor, or set of sensors that accepts input from the user based on haptic and/or tactile contact. Touch screen **112** and display controller **156** (along with any associated modules and/or sets of instructions in memory **102**) detect contact (and any movement or breaking of the contact) on touch screen **112** and convert the detected contact into interaction with user-interface objects (e.g., one or more soft keys, icons, web pages, or images) that are displayed on touch screen **112**. In an exemplary embodiment, a point of contact between touch screen **112** and the user corresponds to a finger of the user.

[0062] Touch screen **112** optionally uses LCD (liquid crystal display) technology, LPD (light emitting polymer display) technology, or LED (light emitting diode) technology, although other display technologies are used in other embodiments. Touch screen **112** and display controller **156** optionally detect contact and any movement or breaking thereof using any of a plurality of touch sensing technologies now known or later developed, including but not limited to capacitive, resistive, infrared, and surface acoustic wave technologies, as well as other proximity sensor arrays or other elements for determining one or more points of contact with touch screen **112**. In an exemplary embodiment, projected mutual capacitance sensing technology is used, such as that found in the iPhone® and iPod Touch® from Apple Inc. of Cupertino, Calif.

[0063] A touch-sensitive display in some embodiments of touch screen **112** is, optionally, analogous to the multi-touch sensitive touchpads described in the following U.S. Pat. No. 6,323,846 (Westerman et al.), U.S. Pat. No. 6,570,557 (Westerman et al.), and/or U.S. Pat. No. 6,677,932 (Westerman), and/or U.S. Patent Publication 2002/0015024A1, each of which is hereby incorporated by reference in its entirety. However, touch screen **112** displays visual output from device **100**, whereas touch-sensitive touchpads do not provide visual output.

[0064] A touch-sensitive display in some embodiments of touch screen **112** is described in the following applications: (1) U.S. patent application Ser. No. 11/381,313, "Multipoint Touch Surface Controller," filed May 2, 2006; (2) U.S. patent application Ser. No. 10/840,862, "Multipoint Touchscreen," filed May 6, 2004; (3) U.S. patent application Ser. No. 10/903,964, "Gestures For Touch Sensitive Input Devices," filed Jul. 30, 2004; (4) U.S. patent application Ser. No. 11/048,264, "Gestures For Touch Sensitive Input Devices," filed Jan. 31, 2005; (5) U.S. patent application Ser. No. 11/038,590, "Mode-Based Graphical User Interfaces For Touch Sensitive Input Devices," filed Jan. 18, 2005; (6) U.S. patent application Ser. No. 11/228,758, "Virtual Input Device Placement On A Touch Screen User Interface," filed Sep. 16, 2005; (7) U.S. patent application Ser. No. 11/228,700, "Operation Of A Computer With A Touch Screen Interface," filed Sep. 16, 2005; (8) U.S. patent application Ser. No. 11/228,737, "Activating Virtual Keys Of A Touch-Screen Virtual Keyboard," filed Sep. 16, 2005; and (9) U.S. patent application Ser. No. 11/367,749, "Multi-Functional Hand-Held Device," filed Mar. 3, 2006. All of these applications are incorporated by reference herein in their entirety.

[0065] Touch screen **112** optionally has a video resolution in excess of 100 dpi. In some embodiments, the touch screen has a video resolution of approximately 160 dpi. The user optionally makes contact with touch screen **112** using any suitable object or appendage, such as a stylus, a finger, and so forth. In some embodiments, the user interface is designed to work primarily with finger-based contacts and gestures, which can be less precise than stylus-based input due to the larger area of contact of a finger on the touch screen. In some embodiments, the device translates the rough finger-based input into a precise pointer/cursor position or command for performing the actions desired by the user.

[0066] In some embodiments, in addition to the touch screen, device **100** optionally includes a touchpad for activating or deactivating particular functions. In some embodiments, the touchpad is a touch-sensitive area of the device that, unlike the touch screen, does not display visual output. The touchpad is, optionally, a touch-sensitive surface that is separate from touch screen **112** or an extension of the touch-sensitive surface formed by the touch screen.

[0067] Device **100** also includes power system **162** for powering the various components. Power system **162** optionally includes a power management system, one or more power sources (e.g., battery, alternating current (AC)), a recharging system, a power failure detection circuit, a power converter or inverter, a power status indicator (e.g., a

light-emitting diode (LED)) and any other components associated with the generation, management and distribution of power in portable devices.

[0068] Device **100** optionally also includes one or more optical sensors **164**. FIG. **1**A shows an optical sensor coupled to optical sensor controller **158** in I/O subsystem **106**. Optical sensor **164** optionally includes charge-coupled device (CCD) or complementary metal-oxide semiconductor (CMOS) phototransistors. Optical sensor **164** receives light from the environment, projected through one or more lenses, and converts the light to data representing an image. In conjunction with imaging module **143** (also called a camera module), optical sensor **164** optionally captures still images or video. In some embodiments, an optical sensor is located on the back of device **100**, opposite touch screen display **112** on the front of the device so that the touch screen display is enabled for use as a viewfinder for still and/or video image acquisition. In some embodiments, an optical sensor is located on the front of the device so that the user's image is, optionally, obtained for video conferencing while the user views the other video conference participants on the touch screen display. In some embodiments, the position of optical sensor **164** can be changed by the user (e.g., by rotating the lens and the sensor in the device housing) so that a single optical sensor **164** is used along with the touch screen display for both video conferencing and still and/or video image acquisition.

[0069] Device **100** optionally also includes one or more depth camera sensors **175**. FIG. **1**A shows a depth camera sensor coupled to depth camera controller **169** in I/O subsystem **106**. Depth camera sensor **175** receives data from the environment to create a three dimensional model of an object (e.g., a face) within a scene from a viewpoint (e.g., a depth camera sensor). In some embodiments, in conjunction with imaging module **143** (also called a camera module), depth camera sensor **175** is optionally used to determine a depth map of different portions of an image captured by the imaging module **143**. In some embodiments, a depth camera sensor is located on the front of device **100** so that the user's image with depth information is, optionally, obtained for video conferencing while the user views the other video conference participants on the touch screen display and to capture selfies with depth map data. In some embodiments, the depth camera sensor **175** is located on the back of device, or on the back and the front of the device **100**. In some embodiments, the position of depth camera sensor **175** can be changed by the user (e.g., by rotating the lens and the sensor in the device housing) so that a depth camera sensor **175** is used along with the touch screen display for both video conferencing and still and/or video image acquisition.

[0070] In some embodiments, a depth map (e.g., depth map image) contains information (e.g., values) that relates to the distance of objects in a scene from a viewpoint (e.g., a camera, an optical sensor, a depth camera sensor). In one embodiment of a depth map, each depth pixel defines the position in the viewpoint's Z-axis where its corresponding two-dimensional pixel is located. In some embodiments, a depth map is composed of pixels wherein each pixel is defined by a value (e.g., 0-255). For example, the "0" value represents pixels that are located at the most distant place in a "three dimensional" scene and the "255" value represents pixels that are located closest to a viewpoint (e.g., a camera, an optical sensor, a depth camera sensor) in the "three dimensional" scene. In other embodiments, a depth map represents the distance between an object in a scene and the plane of the viewpoint. In some embodiments, the depth map includes information about the relative depth of various features of an object of interest in view of the depth camera (e.g., the relative depth of eyes, nose, mouth, ears of a user's face). In some embodiments, the depth map includes information that enables the device to determine contours of the object of interest in a z direction.

[0071] Device **100** optionally also includes one or more contact intensity sensors **165**. FIG. **1**A shows a contact intensity sensor coupled to intensity sensor controller **159** in I/O subsystem **106**. Contact intensity sensor **165** optionally includes one or more piezoresistive strain gauges, capacitive force sensors, electric force sensors, piezoelectric force sensors, optical force sensors, capacitive touch-sensitive surfaces, or other intensity sensors (e.g., sensors used to measure the force (or pressure) of a contact on a touch-sensitive surface). Contact intensity sensor **165** receives contact intensity information (e.g., pressure information or a proxy for pressure information) from the environment. In some embodiments, at least one contact intensity sensor is collocated with, or proximate to, a touch-sensitive surface (e.g., touch-sensitive display system **112**). In some embodiments, at least one contact intensity sensor is located on the back of device **100**, opposite touch screen display **112**, which is located on the front of device **100**.

[0072] Device **100** optionally also includes one or more proximity sensors **166**. FIG. **1**A shows proximity sensor **166** coupled to peripherals interface **118**. Alternately, proximity sensor **166** is, optionally, coupled to input controller **160** in I/O subsystem **106**. Proximity sensor **166** optionally performs as described in U.S. patent application Ser. No. 11/241,839, "Proximity Detector In Handheld Device"; Ser. No. 11/240,788, "Proximity Detector In Handheld Device"; Ser. No. 11/620,702, "Using Ambient Light Sensor To Augment Proximity Sensor Output"; Ser. No. 11/586,862, "Automated Response To And Sensing Of User Activity In Portable Devices"; and Ser. No. 11/638,251, "Methods And Systems For Automatic Configuration Of Peripherals," which are hereby incorporated by reference in their entirety. In some embodiments, the proximity sensor turns off and disables touch screen **112** when the multifunction device is placed near the user's ear (e.g., when the user is making a phone call).

[0073] Device **100** optionally also includes one or more tactile output generators **167**. FIG. **1**A shows a tactile output generator coupled to haptic feedback controller **161** in I/O subsystem **106**. Tactile output generator **167** optionally includes one or more electroacoustic devices such as speakers or other audio components and/or electromechanical devices that convert energy into linear motion such as a motor, solenoid, electroactive polymer, piezoelectric actuator, electrostatic actuator, or other tactile output generating component (e.g., a component that converts electrical signals into tactile outputs on the device). Contact intensity sensor **165** receives tactile feedback generation instructions from haptic feedback module **133** and generates tactile outputs on device **100** that are capable of being sensed by a user of device **100**. In some embodiments, at least one tactile output generator is collocated with, or proximate to, a touch-sensitive surface (e.g., touch-sensitive display system **112**) and, optionally, generates a tactile output by moving the touch-sensitive surface vertically (e.g., in/out of a surface of device **100**) or laterally (e.g., back and forth in the

same plane as a surface of device **100**). In some embodiments, at least one tactile output generator sensor is located on the back of device **100**, opposite touch screen display **112**, which is located on the front of device **100**.

[0074] Device **100** optionally also includes one or more accelerometers **168**. FIG. **1**A shows accelerometer **168** coupled to peripherals interface **118**. Alternately, accelerometer **168** is, optionally, coupled to an input controller **160** in I/O subsystem **106**. Accelerometer **168** optionally performs as described in U.S. Patent Publication No. 20050190059, "Acceleration-based Theft Detection System for Portable Electronic Devices," and U.S. Patent Publication No. 20060017692, "Methods And Apparatuses For Operating A Portable Device Based On An Accelerometer," both of which are incorporated by reference herein in their entirety. In some embodiments, information is displayed on the touch screen display in a portrait view or a landscape view based on an analysis of data received from the one or more accelerometers. Device **100** optionally includes, in addition to accelerometer(s) **168**, a magnetometer and a GPS (or GLONASS or other global navigation system) receiver for obtaining information concerning the location and orientation (e.g., portrait or landscape) of device **100**.

[0075] In some embodiments, the software components stored in memory **102** include operating system **126**, communication module (or set of instructions) **128**, contact/motion module (or set of instructions) **130**, graphics module (or set of instructions) **132**, text input module (or set of instructions) **134**, Global Positioning System (GPS) module (or set of instructions) **135**, and applications (or sets of instructions) **136**. Furthermore, in some embodiments, memory **102** (FIG. **1**A) or **370** (FIG. **3**) stores device/global internal state **157**, as shown in FIGS. **1**A and **3**. Device/global internal state **157** includes one or more of: active application state, indicating which applications, if any, are currently active; display state, indicating what applications, views or other information occupy various regions of touch screen display **112**; sensor state, including information obtained from the device's various sensors and input control devices **116**; and location information concerning the device's location and/or attitude.

[0076] Operating system **126** (e.g., Darwin, RTXC, LINUX, UNIX, OS X, iOS, WINDOWS, or an embedded operating system such as VxWorks) includes various software components and/or drivers for controlling and managing general system tasks (e.g., memory management, storage device control, power management, etc.) and facilitates communication between various hardware and software components.

[0077] Communication module **128** facilitates communication with other devices over one or more external ports **124** and also includes various software components for handling data received by RF circuitry **108** and/or external port **124**. External port **124** (e.g., Universal Serial Bus (USB), FIREWIRE, etc.) is adapted for coupling directly to other devices or indirectly over a network (e.g., the Internet, wireless LAN, etc.). In some embodiments, the external port is a multi-pin (e.g., 30-pin) connector that is the same as, or similar to and/or compatible with, the 30-pin connector used on iPod® (trademark of Apple Inc.) devices.

[0078] Contact/motion module **130** optionally detects contact with touch screen **112** (in conjunction with display controller **156**) and other touch-sensitive devices (e.g., a touchpad or physical click wheel). Contact/motion module **130** includes various software components for performing various operations related to detection of contact, such as determining if contact has occurred (e.g., detecting a finger-down event), determining an intensity of the contact (e.g., the force or pressure of the contact or a substitute for the force or pressure of the contact), determining if there is movement of the contact and tracking the movement across the touch-sensitive surface (e.g., detecting one or more finger-dragging events), and determining if the contact has ceased (e.g., detecting a finger-up event or a break in contact). Contact/motion module **130** receives contact data from the touch-sensitive surface. Determining movement of the point of contact, which is represented by a series of contact data, optionally includes determining speed (magnitude), velocity (magnitude and direction), and/or an acceleration (a change in magnitude and/or direction) of the point of contact. These operations are, optionally, applied to single contacts (e.g., one finger contacts) or to multiple simultaneous contacts (e.g., "multi-touch"/multiple finger contacts). In some embodiments, contact/motion module **130** and display controller **156** detect contact on a touchpad.

[0079] In some embodiments, contact/motion module **130** uses a set of one or more intensity thresholds to determine whether an operation has been performed by a user (e.g., to determine whether a user has "clicked" on an icon). In some embodiments, at least a subset of the intensity thresholds are determined in accordance with software parameters (e.g., the intensity thresholds are not determined by the activation thresholds of particular physical actuators and can be adjusted without changing the physical hardware of device **100**). For example, a mouse "click" threshold of a trackpad or touch screen display can be set to any of a large range of predefined threshold values without changing the trackpad or touch screen display hardware. Additionally, in some implementations, a user of the device is provided with software settings for adjusting one or more of the set of intensity thresholds (e.g., by adjusting individual intensity thresholds and/or by adjusting a plurality of intensity thresholds at once with a system-level click "intensity" parameter).

[0080] Contact/motion module **130** optionally detects a gesture input by a user. Different gestures on the touch-sensitive surface have different contact patterns (e.g., different motions, timings, and/or intensities of detected contacts). Thus, a gesture is, optionally, detected by detecting a particular contact pattern. For example, detecting a finger tap gesture includes detecting a finger-down event followed by detecting a finger-up (liftoff) event at the same position (or substantially the same position) as the finger-down event (e.g., at the position of an icon). As another example, detecting a finger swipe gesture on the touch-sensitive surface includes detecting a finger-down event followed by detecting one or more finger-dragging events, and subsequently followed by detecting a finger-up (liftoff) event.

[0081] Graphics module **132** includes various known software components for rendering and displaying graphics on touch screen **112** or other display, including components for changing the visual impact (e.g., brightness, transparency, saturation, contrast, or other visual property) of graphics that are displayed. As used herein, the term "graphics" includes any object that can be displayed to a user, including, without limitation, text, web pages, icons (such as user-interface objects including soft keys), digital images, videos, animations, and the like.

[0082] In some embodiments, graphics module **132** stores data representing graphics to be used. Each graphic is, optionally, assigned a corresponding code. Graphics module **132** receives, from applications etc., one or more codes specifying graphics to be displayed along with, if necessary, coordinate data and other graphic property data, and then generates screen image data to output to display controller **156**.

[0083] Haptic feedback module **133** includes various software components for generating instructions used by tactile output generator(s) **167** to produce tactile outputs at one or more locations on device **100** in response to user interactions with device **100**.

[0084] Text input module **134**, which is, optionally, a component of graphics module **132**, provides soft keyboards for entering text in various applications (e.g., contacts **137**, e-mail **140**, IM **141**, browser **147**, and any other application that needs text input).

[0085] GPS module **135** determines the location of the device and provides this information for use in various applications (e.g., to telephone **138** for use in location-based dialing; to camera **143** as picture/video metadata; and to applications that provide location-based services such as weather widgets, local yellow page widgets, and map/navigation widgets).

[0086] Applications **136** optionally include the following modules (or sets of instructions), or a subset or superset thereof:

- **[0087]** Contacts module **137** (sometimes called an address book or contact list);
- **[0088]** Telephone module **138**;
- **[0089]** Video conference module **139**;
- **[0090]** E-mail client module **140**;
- **[0091]** Instant messaging (IM) module **141**;
- **[0092]** Workout support module **142**;
- **[0093]** Camera module **143** for still and/or video images;
- **[0094]** Image management module **144**;
- **[0095]** Video player module;
- **[0096]** Music player module;
- **[0097]** Browser module **147**;
- **[0098]** Calendar module **148**;
- **[0099]** Widget modules **149**, which optionally include one or more of: weather widget **149-1**, stocks widget **149-2**, calculator widget **149-3**, alarm clock widget **149-4**, dictionary widget **149-5**, and other widgets obtained by the user, as well as user-created widgets **149-6**;
- **[0100]** Widget creator module **150** for making user-created widgets **149-6**;
- **[0101]** Search module **151**;
- **[0102]** Video and music player module **152**, which merges video player module and music player module;
- **[0103]** Notes module **153**;
- **[0104]** Map module **154**; and/or
- **[0105]** Online video module **155**.

[0106] Examples of other applications **136** that are, optionally, stored in memory **102** include other word processing applications, other image editing applications, drawing applications, presentation applications, JAVA-enabled applications, encryption, digital rights management, voice recognition, and voice replication.

[0107] In conjunction with touch screen **112**, display controller **156**, contact/motion module **130**, graphics module **132**, and text input module **134**, contacts module **137** are, optionally, used to manage an address book or contact list (e.g., stored in application internal state **192** of contacts module **137** in memory **102** or memory **370**), including: adding name(s) to the address book; deleting name(s) from the address book; associating telephone number(s), e-mail address(es), physical address(es) or other information with a name; associating an image with a name; categorizing and sorting names; providing telephone numbers or e-mail addresses to initiate and/or facilitate communications by telephone **138**, video conference module **139**, e-mail **140**, or IM **141**; and so forth.

[0108] In conjunction with RF circuitry **108**, audio circuitry **110**, speaker **111**, microphone **113**, touch screen **112**, display controller **156**, contact/motion module **130**, graphics module **132**, and text input module **134**, telephone module **138** are optionally, used to enter a sequence of characters corresponding to a telephone number, access one or more telephone numbers in contacts module **137**, modify a telephone number that has been entered, dial a respective telephone number, conduct a conversation, and disconnect or hang up when the conversation is completed. As noted above, the wireless communication optionally uses any of a plurality of communications standards, protocols, and technologies.

[0109] In conjunction with RF circuitry **108**, audio circuitry **110**, speaker **111**, microphone **113**, touch screen **112**, display controller **156**, optical sensor **164**, optical sensor controller **158**, contact/motion module **130**, graphics module **132**, text input module **134**, contacts module **137**, and telephone module **138**, video conference module **139** includes executable instructions to initiate, conduct, and terminate a video conference between a user and one or more other participants in accordance with user instructions.

[0110] In conjunction with RF circuitry **108**, touch screen **112**, display controller **156**, contact/motion module **130**, graphics module **132**, and text input module **134**, e-mail client module **140** includes executable instructions to create, send, receive, and manage e-mail in response to user instructions. In conjunction with image management module **144**, e-mail client module **140** makes it very easy to create and send e-mails with still or video images taken with camera module **143**.

[0111] In conjunction with RF circuitry **108**, touch screen **112**, display controller **156**, contact/motion module **130**, graphics module **132**, and text input module **134**, the instant messaging module **141** includes executable instructions to enter a sequence of characters corresponding to an instant message, to modify previously entered characters, to transmit a respective instant message (for example, using a Short Message Service (SMS) or Multimedia Message Service (MMS) protocol for telephony-based instant messages or using XMPP, SIMPLE, or IMPS for Internet-based instant messages), to receive instant messages, and to view received instant messages. In some embodiments, transmitted and/or received instant messages optionally include graphics, photos, audio files, video files and/or other attachments as are supported in an MMS and/or an Enhanced Messaging Service (EMS). As used herein, "instant messaging" refers to both telephony-based messages (e.g., messages sent using SMS or MMS) and Internet-based messages (e.g., messages sent using XMPP, SIMPLE, or IMPS).

[0112] In conjunction with RF circuitry **108**, touch screen **112**, display controller **156**, contact/motion module **130**,

graphics module **132**, text input module **134**, GPS module **135**, map module **154**, and music player module, workout support module **142** includes executable instructions to create workouts (e.g., with time, distance, and/or calorie burning goals); communicate with workout sensors (sports devices); receive workout sensor data; calibrate sensors used to monitor a workout; select and play music for a workout; and display, store, and transmit workout data.

[0113] In conjunction with touch screen **112**, display controller **156**, optical sensor(s) **164**, optical sensor controller **158**, contact/motion module **130**, graphics module **132**, and image management module **144**, camera module **143** includes executable instructions to capture still images or video (including a video stream) and store them into memory **102**, modify characteristics of a still image or video, or delete a still image or video from memory **102**.

[0114] In conjunction with touch screen **112**, display controller **156**, contact/motion module **130**, graphics module **132**, text input module **134**, and camera module **143**, image management module **144** includes executable instructions to arrange, modify (e.g., edit), or otherwise manipulate, label, delete, present (e.g., in a digital slide show or album), and store still and/or video images.

[0115] In conjunction with RF circuitry **108**, touch screen **112**, display controller **156**, contact/motion module **130**, graphics module **132**, and text input module **134**, browser module **147** includes executable instructions to browse the Internet in accordance with user instructions, including searching, linking to, receiving, and displaying web pages or portions thereof, as well as attachments and other files linked to web pages.

[0116] In conjunction with RF circuitry **108**, touch screen **112**, display controller **156**, contact/motion module **130**, graphics module **132**, text input module **134**, e-mail client module **140**, and browser module **147**, calendar module **148** includes executable instructions to create, display, modify, and store calendars and data associated with calendars (e.g., calendar entries, to-do lists, etc.) in accordance with user instructions.

[0117] In conjunction with RF circuitry **108**, touch screen **112**, display controller **156**, contact/motion module **130**, graphics module **132**, text input module **134**, and browser module **147**, widget modules **149** are mini-applications that are, optionally, downloaded and used by a user (e.g., weather widget **149-1**, stocks widget **149-2**, calculator widget **149-3**, alarm clock widget **149-4**, and dictionary widget **149-5**) or created by the user (e.g., user-created widget **149-6**). In some embodiments, a widget includes an HTML (Hypertext Markup Language) file, a CSS (Cascading Style Sheets) file, and a JavaScript file. In some embodiments, a widget includes an XML (Extensible Markup Language) file and a JavaScript file (e.g., Yahoo! Widgets).

[0118] In conjunction with RF circuitry **108**, touch screen **112**, display controller **156**, contact/motion module **130**, graphics module **132**, text input module **134**, and browser module **147**, the widget creator module **150** are, optionally, used by a user to create widgets (e.g., turning a user-specified portion of a web page into a widget).

[0119] In conjunction with touch screen **112**, display controller **156**, contact/motion module **130**, graphics module **132**, and text input module **134**, search module **151** includes executable instructions to search for text, music, sound, image, video, and/or other files in memory **102** that match one or more search criteria (e.g., one or more user-specified search terms) in accordance with user instructions.

[0120] In conjunction with touch screen **112**, display controller **156**, contact/motion module **130**, graphics module **132**, audio circuitry **110**, speaker **111**, RF circuitry **108**, and browser module **147**, video and music player module **152** includes executable instructions that allow the user to download and play back recorded music and other sound files stored in one or more file formats, such as MP3 or AAC files, and executable instructions to display, present, or otherwise play back videos (e.g., on touch screen **112** or on an external, connected display via external port **124**). In some embodiments, device **100** optionally includes the functionality of an MP3 player, such as an iPod (trademark of Apple Inc.).

[0121] In conjunction with touch screen **112**, display controller **156**, contact/motion module **130**, graphics module **132**, and text input module **134**, notes module **153** includes executable instructions to create and manage notes, to-do lists, and the like in accordance with user instructions.

[0122] In conjunction with RF circuitry **108**, touch screen **112**, display controller **156**, contact/motion module **130**, graphics module **132**, text input module **134**, GPS module **135**, and browser module **147**, map module **154** are, optionally, used to receive, display, modify, and store maps and data associated with maps (e.g., driving directions, data on stores and other points of interest at or near a particular location, and other location-based data) in accordance with user instructions.

[0123] In conjunction with touch screen **112**, display controller **156**, contact/motion module **130**, graphics module **132**, audio circuitry **110**, speaker **111**, RF circuitry **108**, text input module **134**, e-mail client module **140**, and browser module **147**, online video module **155** includes instructions that allow the user to access, browse, receive (e.g., by streaming and/or download), play back (e.g., on the touch screen or on an external, connected display via external port **124**), send an e-mail with a link to a particular online video, and otherwise manage online videos in one or more file formats, such as H.264. In some embodiments, instant messaging module **141**, rather than e-mail client module **140**, is used to send a link to a particular online video. Additional description of the online video application can be found in U.S. Provisional Patent Application No. 60/936,562, "Portable Multifunction Device, Method, and Graphical User Interface for Playing Online Videos," filed Jun. 20, 2007, and U.S. patent application Ser. No. 11/968,067, "Portable Multifunction Device, Method, and Graphical User Interface for Playing Online Videos," filed Dec. 31, 2007, the contents of which are hereby incorporated by reference in their entirety.

[0124] Each of the above-identified modules and applications corresponds to a set of executable instructions for performing one or more functions described above and the methods described in this application (e.g., the computer-implemented methods and other information processing methods described herein). These modules (e.g., sets of instructions) need not be implemented as separate software programs, procedures, or modules, and thus various subsets of these modules are, optionally, combined or otherwise rearranged in various embodiments. For example, video player module is, optionally, combined with music player module into a single module (e.g., video and music player module **152**, FIG. **1**A). In some embodiments, memory **102** optionally stores a subset of the modules and data structures

identified above. Furthermore, memory **102** optionally stores additional modules and data structures not described above.

[0125] In some embodiments, device **100** is a device where operation of a predefined set of functions on the device is performed exclusively through a touch screen and/or a touchpad. By using a touch screen and/or a touchpad as the primary input control device for operation of device **100**, the number of physical input control devices (such as push buttons, dials, and the like) on device **100** is, optionally, reduced.

[0126] The predefined set of functions that are performed exclusively through a touch screen and/or a touchpad optionally include navigation between user interfaces. In some embodiments, the touchpad, when touched by the user, navigates device **100** to a main, home, or root menu from any user interface that is displayed on device **100**. In such embodiments, a "menu button" is implemented using a touchpad. In some other embodiments, the menu button is a physical push button or other physical input control device instead of a touchpad.

[0127] FIG. **1**B is a block diagram illustrating exemplary components for event handling in accordance with some embodiments. In some embodiments, memory **102** (FIG. **1**A) or **370** (FIG. **3**) includes event sorter **170** (e.g., in operating system **126**) and a respective application **136-1** (e.g., any of the aforementioned applications **137-151**, **155**, **380-390**).

[0128] Event sorter **170** receives event information and determines the application **136-1** and application view **191** of application **136-1** to which to deliver the event information. Event sorter **170** includes event monitor **171** and event dispatcher module **174**. In some embodiments, application **136-1** includes application internal state **192**, which indicates the current application view(s) displayed on touch-sensitive display **112** when the application is active or executing. In some embodiments, device/global internal state **157** is used by event sorter **170** to determine which application(s) is (are) currently active, and application internal state **192** is used by event sorter **170** to determine application views **191** to which to deliver event information.

[0129] In some embodiments, application internal state **192** includes additional information, such as one or more of: resume information to be used when application **136-1** resumes execution, user interface state information that indicates information being displayed or that is ready for display by application **136-1**, a state queue for enabling the user to go back to a prior state or view of application **136-1**, and a redo/undo queue of previous actions taken by the user.

[0130] Event monitor **171** receives event information from peripherals interface **118**. Event information includes information about a sub-event (e.g., a user touch on touch-sensitive display **112**, as part of a multi-touch gesture). Peripherals interface **118** transmits information it receives from I/O subsystem **106** or a sensor, such as proximity sensor **166**, accelerometer(s) **168**, and/or microphone **113** (through audio circuitry **110**). Information that peripherals interface **118** receives from I/O subsystem **106** includes information from touch-sensitive display **112** or a touch-sensitive surface.

[0131] In some embodiments, event monitor **171** sends requests to the peripherals interface **118** at predetermined intervals. In response, peripherals interface **118** transmits event information. In other embodiments, peripherals interface **118** transmits event information only when there is a significant event (e.g., receiving an input above a predetermined noise threshold and/or for more than a predetermined duration).

[0132] In some embodiments, event sorter **170** also includes a hit view determination module **172** and/or an active event recognizer determination module **173**.

[0133] Hit view determination module **172** provides software procedures for determining where a sub-event has taken place within one or more views when touch-sensitive display **112** displays more than one view. Views are made up of controls and other elements that a user can see on the display.

[0134] Another aspect of the user interface associated with an application is a set of views, sometimes herein called application views or user interface windows, in which information is displayed and touch-based gestures occur. The application views (of a respective application) in which a touch is detected optionally correspond to programmatic levels within a programmatic or view hierarchy of the application. For example, the lowest level view in which a touch is detected is, optionally, called the hit view, and the set of events that are recognized as proper inputs are, optionally, determined based, at least in part, on the hit view of the initial touch that begins a touch-based gesture.

[0135] Hit view determination module **172** receives information related to sub-events of a touch-based gesture. When an application has multiple views organized in a hierarchy, hit view determination module **172** identifies a hit view as the lowest view in the hierarchy which should handle the sub-event. In most circumstances, the hit view is the lowest level view in which an initiating sub-event occurs (e.g., the first sub-event in the sequence of sub-events that form an event or potential event). Once the hit view is identified by the hit view determination module **172**, the hit view typically receives all sub-events related to the same touch or input source for which it was identified as the hit view.

[0136] Active event recognizer determination module **173** determines which view or views within a view hierarchy should receive a particular sequence of sub-events. In some embodiments, active event recognizer determination module **173** determines that only the hit view should receive a particular sequence of sub-events. In other embodiments, active event recognizer determination module **173** determines that all views that include the physical location of a sub-event are actively involved views, and therefore determines that all actively involved views should receive a particular sequence of sub-events. In other embodiments, even if touch sub-events were entirely confined to the area associated with one particular view, views higher in the hierarchy would still remain as actively involved views.

[0137] Event dispatcher module **174** dispatches the event information to an event recognizer (e.g., event recognizer **180**). In embodiments including active event recognizer determination module **173**, event dispatcher module **174** delivers the event information to an event recognizer determined by active event recognizer determination module **173**. In some embodiments, event dispatcher module **174** stores in an event queue the event information, which is retrieved by a respective event receiver **182**.

[0138] In some embodiments, operating system **126** includes event sorter **170**. Alternatively, application **136-1** includes event sorter **170**. In yet other embodiments, event

sorter **170** is a stand-alone module, or a part of another module stored in memory **102**, such as contact/motion module **130**.

[0139] In some embodiments, application **136-1** includes a plurality of event handlers **190** and one or more application views **191**, each of which includes instructions for handling touch events that occur within a respective view of the application's user interface. Each application view **191** of the application **136-1** includes one or more event recognizers **180**. Typically, a respective application view **191** includes a plurality of event recognizers **180**. In other embodiments, one or more of event recognizers **180** are part of a separate module, such as a user interface kit or a higher level object from which application **136-1** inherits methods and other properties. In some embodiments, a respective event handler **190** includes one or more of: data updater **176**, object updater **177**, GUI updater **178**, and/or event data **179** received from event sorter **170**. Event handler **190** optionally utilizes or calls data updater **176**, object updater **177**, or GUI updater **178** to update the application internal state **192**. Alternatively, one or more of the application views **191** include one or more respective event handlers **190**. Also, in some embodiments, one or more of data updater **176**, object updater **177**, and GUI updater **178** are included in a respective application view **191**.

[0140] A respective event recognizer **180** receives event information (e.g., event data **179**) from event sorter **170** and identifies an event from the event information. Event recognizer **180** includes event receiver **182** and event comparator **184**. In some embodiments, event recognizer **180** also includes at least a subset of: metadata **183**, and event delivery instructions **188** (which optionally include sub-event delivery instructions).

[0141] Event receiver **182** receives event information from event sorter **170**. The event information includes information about a sub-event, for example, a touch or a touch movement. Depending on the sub-event, the event information also includes additional information, such as location of the sub-event. When the sub-event concerns motion of a touch, the event information optionally also includes speed and direction of the sub-event. In some embodiments, events include rotation of the device from one orientation to another (e.g., from a portrait orientation to a landscape orientation, or vice versa), and the event information includes corresponding information about the current orientation (also called device attitude) of the device.

[0142] Event comparator **184** compares the event information to predefined event or sub-event definitions and, based on the comparison, determines an event or sub-event, or determines or updates the state of an event or sub-event. In some embodiments, event comparator **184** includes event definitions **186**. Event definitions **186** contain definitions of events (e.g., predefined sequences of sub-events), for example, event 1 (**187-1**), event 2 (**187-2**), and others. In some embodiments, sub-events in an event (**187**) include, for example, touch begin, touch end, touch movement, touch cancellation, and multiple touching. In one example, the definition for event 1 (**187-1**) is a double tap on a displayed object. The double tap, for example, comprises a first touch (touch begin) on the displayed object for a predetermined phase, a first liftoff (touch end) for a predetermined phase, a second touch (touch begin) on the displayed object for a predetermined phase, and a second liftoff (touch end) for a predetermined phase. In another example, the definition for event 2 (**187-2**) is a dragging on a displayed object. The dragging, for example, comprises a touch (or contact) on the displayed object for a predetermined phase, a movement of the touch across touch-sensitive display **112**, and liftoff of the touch (touch end). In some embodiments, the event also includes information for one or more associated event handlers **190**.

[0143] In some embodiments, event definition **187** includes a definition of an event for a respective user-interface object. In some embodiments, event comparator **184** performs a hit test to determine which user-interface object is associated with a sub-event. For example, in an application view in which three user-interface objects are displayed on touch-sensitive display **112**, when a touch is detected on touch-sensitive display **112**, event comparator **184** performs a hit test to determine which of the three user-interface objects is associated with the touch (sub-event). If each displayed object is associated with a respective event handler **190**, the event comparator uses the result of the hit test to determine which event handler **190** should be activated. For example, event comparator **184** selects an event handler associated with the sub-event and the object triggering the hit test.

[0144] In some embodiments, the definition for a respective event (**187**) also includes delayed actions that delay delivery of the event information until after it has been determined whether the sequence of sub-events does or does not correspond to the event recognizer's event type.

[0145] When a respective event recognizer **180** determines that the series of sub-events do not match any of the events in event definitions **186**, the respective event recognizer **180** enters an event impossible, event failed, or event ended state, after which it disregards subsequent sub-events of the touch-based gesture. In this situation, other event recognizers, if any, that remain active for the hit view continue to track and process sub-events of an ongoing touch-based gesture.

[0146] In some embodiments, a respective event recognizer **180** includes metadata **183** with configurable properties, flags, and/or lists that indicate how the event delivery system should perform sub-event delivery to actively involved event recognizers. In some embodiments, metadata **183** includes configurable properties, flags, and/or lists that indicate how event recognizers interact, or are enabled to interact, with one another. In some embodiments, metadata **183** includes configurable properties, flags, and/or lists that indicate whether sub-events are delivered to varying levels in the view or programmatic hierarchy.

[0147] In some embodiments, a respective event recognizer **180** activates event handler **190** associated with an event when one or more particular sub-events of an event are recognized. In some embodiments, a respective event recognizer **180** delivers event information associated with the event to event handler **190**. Activating an event handler **190** is distinct from sending (and deferred sending) sub-events to a respective hit view. In some embodiments, event recognizer **180** throws a flag associated with the recognized event, and event handler **190** associated with the flag catches the flag and performs a predefined process.

[0148] In some embodiments, event delivery instructions **188** include sub-event delivery instructions that deliver event information about a sub-event without activating an event handler. Instead, the sub-event delivery instructions deliver event information to event handlers associated with

the series of sub-events or to actively involved views. Event handlers associated with the series of sub-events or with actively involved views receive the event information and perform a predetermined process.

[0149] In some embodiments, data updater **176** creates and updates data used in application **136-1**. For example, data updater **176** updates the telephone number used in contacts module **137**, or stores a video file used in video player module. In some embodiments, object updater **177** creates and updates objects used in application **136-1**. For example, object updater **177** creates a new user-interface object or updates the position of a user-interface object. GUI updater **178** updates the GUI. For example, GUI updater **178** prepares display information and sends it to graphics module **132** for display on a touch-sensitive display.

[0150] In some embodiments, event handler(s) **190** includes or has access to data updater **176**, object updater **177**, and GUI updater **178**. In some embodiments, data updater **176**, object updater **177**, and GUI updater **178** are included in a single module of a respective application **136-1** or application view **191**. In other embodiments, they are included in two or more software modules.

[0151] It shall be understood that the foregoing discussion regarding event handling of user touches on touch-sensitive displays also applies to other forms of user inputs to operate multifunction devices **100** with input devices, not all of which are initiated on touch screens. For example, mouse movement and mouse button presses, optionally coordinated with single or multiple keyboard presses or holds; contact movements such as taps, drags, scrolls, etc. on touchpads; pen stylus inputs; movement of the device; oral instructions; detected eye movements; biometric inputs; and/or any combination thereof are optionally utilized as inputs corresponding to sub-events which define an event to be recognized.

[0152] FIG. **1**C illustrates a block diagram of system **2100** according to various examples. In some examples, system **2100** implements a digital assistant. The terms "digital assistant," "virtual assistant," "intelligent automated assistant," or "automatic digital assistant" refer to any information processing system that interprets natural language input in spoken and/or textual form to infer user intent, and performs actions based on the inferred user intent. For example, to act on an inferred user intent, the system performs one or more of the following: identifying a task flow with steps and parameters designed to accomplish the inferred user intent, inputting specific requirements from the inferred user intent into the task flow; executing the task flow by invoking programs, methods, services, APIs, or the like; and generating output responses to the user in an audible (e.g., speech) and/or visual form.

[0153] Specifically, a digital assistant is capable of accepting a user request at least partially in the form of a natural language command, request, statement, narrative, and/or inquiry. Typically, the user request seeks either an informational answer or performance of a task by the digital assistant. A satisfactory response to the user request includes a provision of the requested informational answer, a performance of the requested task, or a combination of the two. For example, a user asks the digital assistant a question, such as "Where am I right now?" Based on the user's current location, the digital assistant answers, "You are in Central Park near the west gate." The user also requests the performance of a task, for example, "Please invite my friends to my girlfriend's birthday party next week." In response, the digital assistant can acknowledge the request by saying "Yes, right away," and then send a suitable calendar invite on behalf of the user to each of the user's friends listed in the user's electronic address book. During performance of a requested task, the digital assistant sometimes interacts with the user in a continuous dialogue involving multiple exchanges of information over an extended period of time. There are numerous other ways of interacting with a digital assistant to request information or performance of various tasks. In addition to providing verbal responses and taking programmed actions, the digital assistant also provides responses in other visual or audio forms, e.g., as text, alerts, music, videos, animations, etc.

[0154] As shown in FIG. **1**C, in some examples, a digital assistant is implemented according to a client-server model. The digital assistant includes client-side portion **2102** (hereafter "DA client **2102**") executed on user device **104** and server-side portion **2106** (hereafter "DA server **2106**") executed on server system **2108**. DA client **2102** communicates with DA server **2106** through one or more networks **2110**. DA client **2102** provides client-side functionalities such as user-facing input and output processing and communication with DA server **2106**. DA server **2106** provides server-side functionalities for any number of DA clients **2102** each residing on a respective user device **2104**.

[0155] In some examples, DA server **2106** includes client-facing I/O interface **2112**, one or more processing modules **2114**, data and models **2116**, and I/O interface to external services **2118**. The client-facing I/O interface **2112** facilitates the client-facing input and output processing for DA server **2106**. One or more processing modules **2114** utilize data and models **2116** to process speech input and determine the user's intent based on natural language input. Further, one or more processing modules **2114** perform task execution based on inferred user intent. In some examples, DA server **2106** communicates with external services **120** through network(s) **2110** for task completion or information acquisition. I/O interface to external services **2118** facilitates such communications.

[0156] User device **2104** can be any suitable electronic device. In some examples, user device **2104** is a portable multifunctional device (e.g., device **100**, described above with reference to FIG. **1**A), a multifunctional device or another electronic device (e.g., device **600**, **800**, **1000**) A portable multifunctional device is, for example, a mobile telephone that also contains other functions, such as PDA and/or music player functions. Specific examples of portable multifunction devices include the Apple Watch®, iPhone®, iPod Touch®, and iPad® devices from Apple Inc. of Cupertino, Calif. Other examples of portable multifunction devices include, without limitation, earphones/headphones, speakers, and laptop or tablet computers. Further, in some examples, user device **2104** is a non-portable multifunctional device. In particular, user device **2104** is a desktop computer, a game console, a speaker, a television, or a television set-top box. In some examples, user device **2104** includes a touch-sensitive surface (e.g., touch screen displays and/or touchpads). Further, user device **2104** optionally includes one or more other physical user-interface devices, such as a physical keyboard, a mouse, and/or a joystick. Various examples of electronic devices, such as multifunctional devices, are described below in greater detail.

[0157] Examples of communication network(s) **2110** include local area networks (LAN) and wide area networks (WAN), e.g., the Internet. Communication network(s) **2110** is implemented using any known network protocol, including various wired or wireless protocols, such as, for example, Ethernet, Universal Serial Bus (USB), FIREWIRE, Global System for Mobile Communications (GSM), Enhanced Data GSM Environment (EDGE), code division multiple access (CDMA), time division multiple access (TDMA), Bluetooth, Wi-Fi, voice over Internet Protocol (VoIP), Wi-MAX, or any other suitable communication protocol.

[0158] Server system **2108** is implemented on one or more standalone data processing apparatus or a distributed network of computers. In some examples, server system **2108** also employs various virtual devices and/or services of third-party service providers (e.g., third-party cloud service providers) to provide the underlying computing resources and/or infrastructure resources of server system **2108**.

[0159] In some examples, user device **2104** communicates with DA server **2106** via second user device **2122**. Second user device **2122** is similar or identical to user device **2104**. User device **2104** is configured to communicatively couple to second user device **2122** via a direct communication connection, such as Bluetooth, NFC, BTLE, or the like, or via a wired or wireless network, such as a local Wi-Fi network. In some examples, second user device **2122** is configured to act as a proxy between user device **2104** and DA server **2106**. For example, DA client **2102** of user device **2104** is configured to transmit information (e.g., a user request received at user device **2104**) to DA server **2106** via second user device **2122**. DA server **2106** processes the information and returns relevant data (e.g., data content responsive to the user request) to user device **2104** via second user device **2122**.

[0160] In some examples, user device **2104** is configured to communicate abbreviated requests for data to second user device **2122** to reduce the amount of information transmitted from user device **2104**. Second user device **2122** is configured to determine supplemental information to add to the abbreviated request to generate a complete request to transmit to DA server **2106**. This system architecture can advantageously allow user device **2104** having limited communication capabilities and/or limited battery power (e.g., a watch or a similar compact electronic device) to access services provided by DA server **2106** by using second user device **2122**, having greater communication capabilities and/or battery power (e.g., a mobile phone, laptop computer, tablet computer, or the like), as a proxy to DA server **2106**. While only two user devices **2104** and **2122** are shown in FIG. **1**C, it should be appreciated that system **2100**, in some examples, includes any number and type of user devices configured in this proxy configuration to communicate with DA server system **2106**.

[0161] Although the digital assistant shown in FIG. **1**C includes both a client-side portion (e.g., DA client **2102**) and a server-side portion (e.g., DA server **2106**), in some examples, the functions of a digital assistant are implemented as a standalone application installed on a user device. In addition, the divisions of functionalities between the client and server portions of the digital assistant can vary in different implementations. For instance, in some examples, the DA client is a thin-client that provides only user-facing input and output processing functions, and delegates all other functionalities of the digital assistant to a backend server.

[0162] FIG. **2** illustrates a portable multifunction device **100** having a touch screen **112** in accordance with some embodiments. The touch screen optionally displays one or more graphics within user interface (UI) **200**. In this embodiment, as well as others described below, a user is enabled to select one or more of the graphics by making a gesture on the graphics, for example, with one or more fingers **202** (not drawn to scale in the figure) or one or more styluses **203** (not drawn to scale in the figure). In some embodiments, selection of one or more graphics occurs when the user breaks contact with the one or more graphics. In some embodiments, the gesture optionally includes one or more taps, one or more swipes (from left to right, right to left, upward and/or downward), and/or a rolling of a finger (from right to left, left to right, upward and/or downward) that has made contact with device **100**. In some implementations or circumstances, inadvertent contact with a graphic does not select the graphic. For example, a swipe gesture that sweeps over an application icon optionally does not select the corresponding application when the gesture corresponding to selection is a tap.

[0163] Device **100** optionally also include one or more physical buttons, such as "home" or menu button **204**. As described previously, menu button **204** is, optionally, used to navigate to any application **136** in a set of applications that are, optionally, executed on device **100**. Alternatively, in some embodiments, the menu button is implemented as a soft key in a GUI displayed on touch screen **112**.

[0164] In some embodiments, device **100** includes touch screen **112**, menu button **204**, push button **206** for powering the device on/off and locking the device, volume adjustment button(s) **208**, subscriber identity module (SIM) card slot **210**, headset jack **212**, and docking/charging external port **124**. Push button **206** is, optionally, used to turn the power on/off on the device by depressing the button and holding the button in the depressed state for a predefined time interval; to lock the device by depressing the button and releasing the button before the predefined time interval has elapsed; and/or to unlock the device or initiate an unlock process. In an alternative embodiment, device **100** also accepts verbal input for activation or deactivation of some functions through microphone **113**. Device **100** also, optionally, includes one or more contact intensity sensors **165** for detecting intensity of contacts on touch screen **112** and/or one or more tactile output generators **167** for generating tactile outputs for a user of device **100**.

[0165] FIG. **3** is a block diagram of an exemplary multifunction device with a display and a touch-sensitive surface in accordance with some embodiments. Device **300** need not be portable. In some embodiments, device **300** is a laptop computer, a desktop computer, a tablet computer, a multimedia player device, a navigation device, an educational device (such as a child's learning toy), a gaming system, or a control device (e.g., a home or industrial controller). Device **300** typically includes one or more processing units (CPUs) **310**, one or more network or other communications interfaces **360**, memory **370**, and one or more communication buses **320** for interconnecting these components. Communication buses **320** optionally include circuitry (sometimes called a chipset) that interconnects and controls communications between system components. Device **300**

includes input/output (I/O) interface **330** comprising display **340**, which is typically a touch screen display. I/O interface **330** also optionally includes a keyboard and/or mouse (or other pointing device) **350** and touchpad **355**, tactile output generator **357** for generating tactile outputs on device **300** (e.g., similar to tactile output generator(s) **167** described above with reference to FIG. **1**A), sensors **359** (e.g., optical, acceleration, proximity, touch-sensitive, and/or contact intensity sensors similar to contact intensity sensor(s) **165** described above with reference to FIG. **1**A). Memory **370** includes high-speed random access memory, such as DRAM, SRAM, DDR RAM, or other random access solid state memory devices; and optionally includes non-volatile memory, such as one or more magnetic disk storage devices, optical disk storage devices, flash memory devices, or other non-volatile solid state storage devices. Memory **370** optionally includes one or more storage devices remotely located from CPU(s) **310**. In some embodiments, memory **370** stores programs, modules, and data structures analogous to the programs, modules, and data structures stored in memory **102** of portable multifunction device **100** (FIG. **1**A), or a subset thereof. Furthermore, memory **370** optionally stores additional programs, modules, and data structures not present in memory **102** of portable multifunction device **100**. For example, memory **370** of device **300** optionally stores drawing module **380**, presentation module **382**, word processing module **384**, website creation module **386**, disk authoring module **388**, and/or spreadsheet module **390**, while memory **102** of portable multifunction device **100** (FIG. **1**A) optionally does not store these modules.

[0166] Each of the above-identified elements in FIG. **3** is, optionally, stored in one or more of the previously mentioned memory devices. Each of the above-identified modules corresponds to a set of instructions for performing a function described above. The above-identified modules or programs (e.g., sets of instructions) need not be implemented as separate software programs, procedures, or modules, and thus various subsets of these modules are, optionally, combined or otherwise rearranged in various embodiments. In some embodiments, memory **370** optionally stores a subset of the modules and data structures identified above. Furthermore, memory **370** optionally stores additional modules and data structures not described above.

[0167] Attention is now directed towards embodiments of user interfaces that are, optionally, implemented on, for example, portable multifunction device **100**.

[0168] FIG. **4**A illustrates an exemplary user interface for a menu of applications on portable multifunction device **100** in accordance with some embodiments. Similar user interfaces are, optionally, implemented on device **300**. In some embodiments, user interface **400** includes the following elements, or a subset or superset thereof:

[0169] Signal strength indicator(s) **402** for wireless communication(s), such as cellular and Wi-Fi signals;

[0170] Time **404**;

[0171] Bluetooth indicator **405**;

[0172] Battery status indicator **406**;

[0173] Tray **408** with icons for frequently used applications, such as:

[0174] Icon **416** for telephone module **138**, labeled "Phone," which optionally includes an indicator **414** of the number of missed calls or voicemail messages;

[0175] Icon **418** for e-mail client module **140**, labeled "Mail," which optionally includes an indicator **410** of the number of unread e-mails;

[0176] Icon **420** for browser module **147**, labeled "Browser;" and

[0177] Icon **422** for video and music player module **152**, also referred to as iPod (trademark of Apple Inc.) module **152**, labeled "iPod;" and

[0178] Icons for other applications, such as:

[0179] Icon **424** for IM module **141**, labeled "Messages;"

[0180] Icon **426** for calendar module **148**, labeled "Calendar;"

[0181] Icon **428** for image management module **144**, labeled "Photos;"

[0182] Icon **430** for camera module **143**, labeled "Camera;"

[0183] Icon **432** for online video module **155**, labeled "Online Video;"

[0184] Icon **434** for stocks widget **149-2**, labeled "Stocks;"

[0185] Icon **436** for map module **154**, labeled "Maps;"

[0186] Icon **438** for weather widget **149-1**, labeled "Weather;"

[0187] Icon **440** for alarm clock widget **149-4**, labeled "Clock;"

[0188] Icon **442** for workout support module **142**, labeled "Workout Support;"

[0189] Icon **444** for notes module **153**, labeled "Notes;" and

[0190] Icon **446** for a settings application or module, labeled "Settings," which provides access to settings for device **100** and its various applications **136**.

[0191] It should be noted that the icon labels illustrated in FIG. **4**A are merely exemplary. For example, icon **422** for video and music player module **152** is labeled "Music" or "Music Player." Other labels are, optionally, used for various application icons. In some embodiments, a label for a respective application icon includes a name of an application corresponding to the respective application icon. In some embodiments, a label for a particular application icon is distinct from a name of an application corresponding to the particular application icon.

[0192] FIG. **4**B illustrates an exemplary user interface on a device (e.g., device **300**, FIG. **3**) with a touch-sensitive surface **451** (e.g., a tablet or touchpad **355**, FIG. **3**) that is separate from the display **450** (e.g., touch screen display **112**). Device **300** also, optionally, includes one or more contact intensity sensors (e.g., one or more of sensors **359**) for detecting intensity of contacts on touch-sensitive surface **451** and/or one or more tactile output generators **357** for generating tactile outputs for a user of device **300**.

[0193] Although some of the examples that follow will be given with reference to inputs on touch screen display **112** (where the touch-sensitive surface and the display are combined), in some embodiments, the device detects inputs on a touch-sensitive surface that is separate from the display, as shown in FIG. **4**B. In some embodiments, the touch-sensitive surface (e.g., **451** in FIG. **4**B) has a primary axis (e.g., **452** in FIG. **4**B) that corresponds to a primary axis (e.g., **453** in FIG. **4**B) on the display (e.g., **450**). In accordance with these embodiments, the device detects contacts (e.g., **460** and **462** in FIG. **4**B) with the touch-sensitive surface **451** at

locations that correspond to respective locations on the display (e.g., in FIG. **4**B, **460** corresponds to **468** and **462** corresponds to **470**). In this way, user inputs (e.g., contacts **460** and **462**, and movements thereof) detected by the device on the touch-sensitive surface (e.g., **451** in FIG. **4**B) are used by the device to manipulate the user interface on the display (e.g., **450** in FIG. **4**B) of the multifunction device when the touch-sensitive surface is separate from the display. It should be understood that similar methods are, optionally, used for other user interfaces described herein.

[0194] Additionally, while the following examples are given primarily with reference to finger inputs (e.g., finger contacts, finger tap gestures, finger swipe gestures), it should be understood that, in some embodiments, one or more of the finger inputs are replaced with input from another input device (e.g., a mouse-based input or stylus input). For example, a swipe gesture is, optionally, replaced with a mouse click (e.g., instead of a contact) followed by movement of the cursor along the path of the swipe (e.g., instead of movement of the contact). As another example, a tap gesture is, optionally, replaced with a mouse click while the cursor is located over the location of the tap gesture (e.g., instead of detection of the contact followed by ceasing to detect the contact). Similarly, when multiple user inputs are simultaneously detected, it should be understood that multiple computer mice are, optionally, used simultaneously, or a mouse and finger contacts are, optionally, used simultaneously.

[0195] FIG. **5**A illustrates exemplary personal electronic device **500**. Device **500** includes body **502**. In some embodiments, device **500** can include some or all of the features described with respect to devices **100** and **300** (e.g., FIGS. **1**A-**4**B). In some embodiments, device **500** has touch-sensitive display screen **504**, hereafter touch screen **504**. Alternatively, or in addition to touch screen **504**, device **500** has a display and a touch-sensitive surface. As with devices **100** and **300**, in some embodiments, touch screen **504** (or the touch-sensitive surface) optionally includes one or more intensity sensors for detecting intensity of contacts (e.g., touches) being applied. The one or more intensity sensors of touch screen **504** (or the touch-sensitive surface) can provide output data that represents the intensity of touches. The user interface of device **500** can respond to touches based on their intensity, meaning that touches of different intensities can invoke different user interface operations on device **500**.

[0196] Exemplary techniques for detecting and processing touch intensity are found, for example, in related applications: International Patent Application Serial No. PCT/US2013/040061, titled "Device, Method, and Graphical User Interface for Displaying User Interface Objects Corresponding to an Application," filed May 8, 2013, published as WIPO Publication No. WO/2013/169849, and International Patent Application Serial No. PCT/US2013/069483, titled "Device, Method, and Graphical User Interface for Transitioning Between Touch Input to Display Output Relationships," filed Nov. 11, 2013, published as WIPO Publication No. WO/2014/105276, each of which is hereby incorporated by reference in their entirety.

[0197] In some embodiments, device **500** has one or more input mechanisms **506** and **508**. Input mechanisms **506** and **508**, if included, can be physical. Examples of physical input mechanisms include push buttons and rotatable mechanisms. In some embodiments, device **500** has one or more attachment mechanisms. Such attachment mechanisms, if included, can permit attachment of device **500** with, for example, hats, eyewear, earrings, necklaces, shirts, jackets, bracelets, watch straps, chains, trousers, belts, shoes, purses, backpacks, and so forth. These attachment mechanisms permit device **500** to be worn by a user.

[0198] FIG. **5**B depicts exemplary personal electronic device **500**. In some embodiments, device **500** can include some or all of the components described with respect to FIGS. **1**A, **1**B, and **3**. Device **500** has bus **512** that operatively couples I/O section **514** with one or more computer processors **516** and memory **518**. I/O section **514** can be connected to display **504**, which can have touch-sensitive component **522** and, optionally, intensity sensor **524** (e.g., contact intensity sensor). In addition, I/O section **514** can be connected with communication unit **530** for receiving application and operating system data, using Wi-Fi, Bluetooth, near field communication (NFC), cellular, and/or other wireless communication techniques. Device **500** can include input mechanisms **506** and/or **508**. Input mechanism **506** is, optionally, a rotatable input device or a depressible and rotatable input device, for example. Input mechanism **508** is, optionally, a button, in some examples.

[0199] Input mechanism **508** is, optionally, a microphone, in some examples. Personal electronic device **500** optionally includes various sensors, such as GPS sensor **532**, accelerometer **534**, directional sensor **540** (e.g., compass), gyroscope **536**, motion sensor **538**, and/or a combination thereof, all of which can be operatively connected to I/O section **514**.

[0200] Memory **518** of personal electronic device **500** can include one or more non-transitory computer-readable storage mediums, for storing computer-executable instructions, which, when executed by one or more computer processors **516**, for example, can cause the computer processors to perform the techniques described below, including processes **700**, **900**, and **1100** (FIGS. **7**A-**7**B, **9**A-**9**B, and **11**A-**11**B). A computer-readable storage medium can be any medium that can tangibly contain or store computer-executable instructions for use by or in connection with the instruction execution system, apparatus, or device. In some examples, the storage medium is a transitory computer-readable storage medium. In some examples, the storage medium is a non-transitory computer-readable storage medium. The non-transitory computer-readable storage medium can include, but is not limited to, magnetic, optical, and/or semiconductor storages. Examples of such storage include magnetic disks, optical discs based on CD, DVD, or Blu-ray technologies, as well as persistent solid-state memory such as flash, solid-state drives, and the like. Personal electronic device **500** is not limited to the components and configuration of FIG. **5**B, but can include other or additional components in multiple configurations.

[0201] As used here, the term "affordance" refers to a user-interactive graphical user interface object that is, optionally, displayed on the display screen of devices **100**, **300**, and/or **500** (FIGS. **1**A, **3**, and **5**A-**5**B). For example, an image (e.g., icon), a button, and text (e.g., hyperlink) each optionally constitute an affordance.

[0202] As used herein, the term "focus selector" refers to an input element that indicates a current part of a user interface with which a user is interacting. In some implementations that include a cursor or other location marker, the cursor acts as a "focus selector" so that when an input (e.g., a press input) is detected on a touch-sensitive surface (e.g., touchpad **355** in FIG. **3** or touch-sensitive surface **451** in

FIG. **4**B) while the cursor is over a particular user interface element (e.g., a button, window, slider, or other user interface element), the particular user interface element is adjusted in accordance with the detected input. In some implementations that include a touch screen display (e.g., touch-sensitive display system **112** in FIG. **1**A or touch screen **112** in FIG. **4**A) that enables direct interaction with user interface elements on the touch screen display, a detected contact on the touch screen acts as a "focus selector" so that when an input (e.g., a press input by the contact) is detected on the touch screen display at a location of a particular user interface element (e.g., a button, window, slider, or other user interface element), the particular user interface element is adjusted in accordance with the detected input. In some implementations, focus is moved from one region of a user interface to another region of the user interface without corresponding movement of a cursor or movement of a contact on a touch screen display (e.g., by using a tab key or arrow keys to move focus from one button to another button); in these implementations, the focus selector moves in accordance with movement of focus between different regions of the user interface. Without regard to the specific form taken by the focus selector, the focus selector is generally the user interface element (or contact on a touch screen display) that is controlled by the user so as to communicate the user's intended interaction with the user interface (e.g., by indicating, to the device, the element of the user interface with which the user is intending to interact). For example, the location of a focus selector (e.g., a cursor, a contact, or a selection box) over a respective button while a press input is detected on the touch-sensitive surface (e.g., a touchpad or touch screen) will indicate that the user is intending to activate the respective button (as opposed to other user interface elements shown on a display of the device).

[0203] As used in the specification and claims, the term "characteristic intensity" of a contact refers to a characteristic of the contact based on one or more intensities of the contact. In some embodiments, the characteristic intensity is based on multiple intensity samples. The characteristic intensity is, optionally, based on a predefined number of intensity samples, or a set of intensity samples collected during a predetermined time period (e.g., 0.05, 0.1, 0.2, 0.5, 1, 2, 5, 10 seconds) relative to a predefined event (e.g., after detecting the contact, prior to detecting liftoff of the contact, before or after detecting a start of movement of the contact, prior to detecting an end of the contact, before or after detecting an increase in intensity of the contact, and/or before or after detecting a decrease in intensity of the contact). A characteristic intensity of a contact is, optionally, based on one or more of: a maximum value of the intensities of the contact, a mean value of the intensities of the contact, an average value of the intensities of the contact, a top 10 percentile value of the intensities of the contact, a value at the half maximum of the intensities of the contact, a value at the 90 percent maximum of the intensities of the contact, or the like. In some embodiments, the duration of the contact is used in determining the characteristic intensity (e.g., when the characteristic intensity is an average of the intensity of the contact over time). In some embodiments, the characteristic intensity is compared to a set of one or more intensity thresholds to determine whether an operation has been performed by a user. For example, the set of one or more intensity thresholds optionally includes a first intensity threshold and a second intensity threshold. In this example, a contact with a characteristic intensity that does not exceed the first threshold results in a first operation, a contact with a characteristic intensity that exceeds the first intensity threshold and does not exceed the second intensity threshold results in a second operation, and a contact with a characteristic intensity that exceeds the second threshold results in a third operation. In some embodiments, a comparison between the characteristic intensity and one or more thresholds is used to determine whether or not to perform one or more operations (e.g., whether to perform a respective operation or forgo performing the respective operation), rather than being used to determine whether to perform a first operation or a second operation.

[0204] In some embodiments, a portion of a gesture is identified for purposes of determining a characteristic intensity. For example, a touch-sensitive surface optionally receives a continuous swipe contact transitioning from a start location and reaching an end location, at which point the intensity of the contact increases. In this example, the characteristic intensity of the contact at the end location is, optionally, based on only a portion of the continuous swipe contact, and not the entire swipe contact (e.g., only the portion of the swipe contact at the end location). In some embodiments, a smoothing algorithm is, optionally, applied to the intensities of the swipe contact prior to determining the characteristic intensity of the contact. For example, the smoothing algorithm optionally includes one or more of: an unweighted sliding-average smoothing algorithm, a triangular smoothing algorithm, a median filter smoothing algorithm, and/or an exponential smoothing algorithm. In some circumstances, these smoothing algorithms eliminate narrow spikes or dips in the intensities of the swipe contact for purposes of determining a characteristic intensity.

[0205] The intensity of a contact on the touch-sensitive surface is, optionally, characterized relative to one or more intensity thresholds, such as a contact-detection intensity threshold, a light press intensity threshold, a deep press intensity threshold, and/or one or more other intensity thresholds. In some embodiments, the light press intensity threshold corresponds to an intensity at which the device will perform operations typically associated with clicking a button of a physical mouse or a trackpad. In some embodiments, the deep press intensity threshold corresponds to an intensity at which the device will perform operations that are different from operations typically associated with clicking a button of a physical mouse or a trackpad. In some embodiments, when a contact is detected with a characteristic intensity below the light press intensity threshold (e.g., and above a nominal contact-detection intensity threshold below which the contact is no longer detected), the device will move a focus selector in accordance with movement of the contact on the touch-sensitive surface without performing an operation associated with the light press intensity threshold or the deep press intensity threshold. Generally, unless otherwise stated, these intensity thresholds are consistent between different sets of user interface figures.

[0206] An increase of characteristic intensity of the contact from an intensity below the light press intensity threshold to an intensity between the light press intensity threshold and the deep press intensity threshold is sometimes referred to as a "light press" input. An increase of characteristic intensity of the contact from an intensity below the deep press intensity threshold to an intensity above the deep press

intensity threshold is sometimes referred to as a "deep press" input. An increase of characteristic intensity of the contact from an intensity below the contact-detection intensity threshold to an intensity between the contact-detection intensity threshold and the light press intensity threshold is sometimes referred to as detecting the contact on the touch-surface. A decrease of characteristic intensity of the contact from an intensity above the contact-detection intensity threshold to an intensity below the contact-detection intensity threshold is sometimes referred to as detecting liftoff of the contact from the touch-surface. In some embodiments, the contact-detection intensity threshold is zero. In some embodiments, the contact-detection intensity threshold is greater than zero.

[0207] In some embodiments described herein, one or more operations are performed in response to detecting a gesture that includes a respective press input or in response to detecting the respective press input performed with a respective contact (or a plurality of contacts), where the respective press input is detected based at least in part on detecting an increase in intensity of the contact (or plurality of contacts) above a press-input intensity threshold. In some embodiments, the respective operation is performed in response to detecting the increase in intensity of the respective contact above the press-input intensity threshold (e.g., a "down stroke" of the respective press input). In some embodiments, the press input includes an increase in intensity of the respective contact above the press-input intensity threshold and a subsequent decrease in intensity of the contact below the press-input intensity threshold, and the respective operation is performed in response to detecting the subsequent decrease in intensity of the respective contact below the press-input threshold (e.g., an "up stroke" of the respective press input).

[0208] In some embodiments, the device employs intensity hysteresis to avoid accidental inputs sometimes termed "jitter," where the device defines or selects a hysteresis intensity threshold with a predefined relationship to the press-input intensity threshold (e.g., the hysteresis intensity threshold is X intensity units lower than the press-input intensity threshold or the hysteresis intensity threshold is 75%, 90%, or some reasonable proportion of the press-input intensity threshold). Thus, in some embodiments, the press input includes an increase in intensity of the respective contact above the press-input intensity threshold and a subsequent decrease in intensity of the contact below the hysteresis intensity threshold that corresponds to the press-input intensity threshold, and the respective operation is performed in response to detecting the subsequent decrease in intensity of the respective contact below the hysteresis intensity threshold (e.g., an "up stroke" of the respective press input). Similarly, in some embodiments, the press input is detected only when the device detects an increase in intensity of the contact from an intensity at or below the hysteresis intensity threshold to an intensity at or above the press-input intensity threshold and, optionally, a subsequent decrease in intensity of the contact to an intensity at or below the hysteresis intensity, and the respective operation is performed in response to detecting the press input (e.g., the increase in intensity of the contact or the decrease in intensity of the contact, depending on the circumstances).

[0209] For ease of explanation, the descriptions of operations performed in response to a press input associated with a press-input intensity threshold or in response to a gesture including the press input are, optionally, triggered in response to detecting either: an increase in intensity of a contact above the press-input intensity threshold, an increase in intensity of a contact from an intensity below the hysteresis intensity threshold to an intensity above the press-input intensity threshold, a decrease in intensity of the contact below the press-input intensity threshold, and/or a decrease in intensity of the contact below the hysteresis intensity threshold corresponding to the press-input intensity threshold. Additionally, in examples where an operation is described as being performed in response to detecting a decrease in intensity of a contact below the press-input intensity threshold, the operation is, optionally, performed in response to detecting a decrease in intensity of the contact below a hysteresis intensity threshold corresponding to, and lower than, the press-input intensity threshold.

[0210] Attention is now directed towards embodiments of user interfaces ("UI") and associated processes that are implemented on an electronic device, such as portable multifunction device **100**, device **300**, or device **500**.

[0211] FIGS. **6**A-**6**O illustrate exemplary user interfaces for activating and interacting with a digital assistant using glance and voice commands, in accordance with some embodiments. These figures are used to illustrate the processes described below, including the processes in FIGS. **7**A-**7**B.

[0212] FIGS. **6**A-**6**D illustrate user **620** activating an exemplary digital assistant on electronic device **600** by glancing at electronic device **600**. Once the digital assistant becomes activated, user **620** turns on the table lamp by speaking a command **624**A-**624**B to the digital assistant.

[0213] FIG. **6**A illustrates a digital assistant on electronic device **600** that is in an inactive state as indicated by display **602** being off. In some embodiments, the digital assistant is in an inactive state when electronic device **600** does not detect a user gaze or user **620** looking at electronic device **600**. In some embodiments, the digital assistant is in an inactive state prior to the electronic device **600** detecting a button press (or other activation command) and/or prior to another device (e.g., watch **630**) having receiving an activation command and transmitting an activation instruction to the electronic device. User **620** may glance at electronic device **600** to trigger the activation of the digital assistant when the gaze information based on the gaze of user **620** satisfies a set of one or more activation criteria. The electronic device **600** obtains gaze information of the gaze of user **620** using one or more camera sensors optionally located on external devices **616**A and/or integrated with electronic device **600**, such as camera **616**B. Camera sensors **616**A are wirelessly connected to electronic device **600** or alternatively could be wired. The one or more camera sensors include infrared camera sensors and/or visible light sensors that measure gaze information once a gaze has been detected. The gaze information optionally includes one or more of: the position of the user's head position, the dwell time or the duration of the gaze, the direction of the gaze, the field of view **606** of the gaze, and whether an object (e.g., electronic device **600** or an external device such as table lamp **618**) is within the field of view **606** of the gaze.

[0214] Electronic device **600** determines whether the gaze information obtained from the one or more camera sensors **616**A-**616**B satisfies a set of one or more activation criteria. In other words, electronic device **600** determines whether user **620** intended to look at electronic device **600** to activate

the digital assistant. The set of one or more activation criteria optionally includes a direction criterion that is satisfied when the user's gaze is directed in the direction of electronic device **600**. The set of one or more activation criteria also optionally includes a dwell time criterion that is satisfied when electronic device **600** detects a dwell time or duration of the gaze in the direction of electronic device **600** for greater than a non-zero, predetermined period of time (e.g., a couple seconds). If the dwell time or duration of the gaze is less than the threshold time period, this may indicate that the user does not intend to trigger the digital assistant. Thus, a dwell time of less than the threshold time period does not satisfy the activation criteria, resulting in the digital assistant not being activated.

[0215] The set of one or more activation criteria also optionally includes a field of view criteria. As illustrated in FIG. **6**A, user **620** is not looking at electronic device **600** and electronic device **600** determines that device **600** is not in the field of view **606** of user **620**. As further discussed below, the field of view criterion is satisfied when electronic device **600** is determined to be in field of view **606** of user **620**, but is not satisfied when the electronic device **600** is determined to not be in the field of view **600** of use **620**. Using the field of view **606** to determine whether a user **602** is looking at a target electronic device (e.g., device **600**) results in a higher gaze detection rate as a gaze may be detected in the field of view even if the user is not looking directly at the target electronic device, but still intends to identify (or activate) the target electronic device.

[0216] The field of view **606** is a range of degrees above and below the user's line-of-sight **612** as indicated by the threshold offset angle **604**. In some embodiments, the field of view criterion is satisfied when electronic device **600** is directly in the line-of-sight **612** of user **620**, or the determined offset degrees is zero. In some embodiments, a field of view criterion is satisfied when a determined offset angle **622** of the gaze is less than the threshold offset angle **604** (e.g., the maximum acceptable deviation from a line-of-sight **612** of user **620** looking directly at the electronic device as illustrated in FIG. **6**B). The determined offset angle **622** (in this example) and the threshold offset angle **604** are greater than zero degrees (e.g., thirty degrees). The determined offset angle **622** is calculated based on the angle formed between a line-of-sight **612** from the gaze of user **620** (e.g., where the user is actually looking) and a calculated line-of-sight **608** from user **620** to electronic device **600** (e.g., the ideal path of the user looking directly at electronic device **600**). When the offset angle **622** is greater than threshold offset angle **604**, the field of view criterion is not satisfied resulting in the digital assistant remaining not activated because electronic device **600** is outside field of view **606**, as illustrated in FIG. **6**A. In some embodiments, if electronic device **600** is determined to be positioned more than a non-zero threshold number of degrees (e.g., more than 30 degrees) outside the field of view **606** of user **620**, electronic device **600** is not in the user's gaze and the digital assistant is not activated. In some embodiments, line-of-sight **612** is measured based on the head position of the user. In some embodiments, the head position of the user is based on the tilt or rotation (e.g., yaw, pitch, and/or roll) of the user's head.

[0217] In FIG. **6**A, the digital assistant on electronic device **600** is not activated because the gaze information obtained from the camera sensors (e.g., **616**A, **61**B) does not satisfy the set of one or more activation criteria, as discussed above. Electronic device **600** determines that user **620** is not looking at electronic device **600** when the dwell time of the gaze of user **620** is less than the threshold dwell time or electronic device **600** is not in the field of view of user **620**. In some embodiments, the field of view criterion is not satisfied when the determined offset angle **622** between the user's line-of-sight **612** and a calculated line-of-sight **608** is greater than the threshold offset angle **604**. When the digital assistant is not activated and the gaze information does not satisfy the set of one or more activation criteria, electronic device **600** forgoes activating the digital assistant. When the digital assistant is not activated, the electronic device **600** optionally does not turn on one or more of a microphone or a speaker, does not record audio, and/or does not enable processing of recorded audio for the purpose of performing a spoken command (e.g., electronic device **600** can still process audio to detect trigger words, but does not process the audio for performance of non-activation commands). In some embodiments, electronic device **600** does not provide any indications using a visual or audio indicator when the digital assistant is not activated (e.g., the state of electronic device **600** remains unchanged). In some embodiments, electronic device **600** remains in the same inactive state that it was in prior to the determination that the gaze information has not satisfied the set of one or more activation criteria.

[0218] In FIG. **6**A, user **620** starts speaking a portion **624**A "Turn . . . " of command **624**A-**624**B "Turn on the table lamp" before the digital assistant is activated. In some embodiments, portion **624**A of command **624**A-**624**B is not processed, is cancelled, and/or is ignored by the digital assistant if the digital assistant is not subsequently activated (e.g., within a set (non-zero) duration of time), as further discussed below. In some embodiments, portion **624**A of command **624**A-**624**B received by electronic device **600** prior to the digital assistant being activated is processed if the digital assistant is activated (e.g., by user **620** looking at electronic device **600**, satisfying the set of one or more activation criteria) prior to completing command **624**A-**624**B (or within the set duration of time), as further discussed below.

[0219] As illustrated in FIG. **6**B, user **620** is looking at electronic device **600** and electronic device **600** determines that electronic device **600** is in the field of view **606** of user **620**. The gaze information, obtained from the camera sensors (e.g., **616**A, **616**B), satisfies the set of one or more activation criteria.

[0220] In FIG. **6**C, electronic device **600** determines that the gaze information, obtained from the camera sensors (e.g., **616**A, **616**B), satisfies the set of one or more activation criteria and, in response, activates the digital assistant. Activating the digital assistant optionally includes one or more of: turning on a microphone or speaker, recording audio, and enabling processing of recorded audio for the purpose of performing a spoken command (rather than merely detecting a trigger phrase). Electronic device **600** optionally provides an indication (e.g., via indicator **610**A, **610**B, or **610**C) that the set of one or more activation criteria has been satisfied and/or that the digital assistant has been activated. The indicator is optionally a visual indication **610**A, an audio indication **610**B, the digital assistant speaking a confirmation **610**C, or any combination of the different indicators. Providing visual indication **610**A optionally includes one or more of: turning on display **602** of electronic

device **600**, turning on an LED of electronic device **600**, displaying a light of a particular color (e.g., blue or a color other than the default color when the digital assistant is not activated), a light pattern, or a visual output on display **602** of electronic device **600**. In some embodiments, visual indicator **610**A indicating that the digital assistant is activated is different from the visual indicator **610**A indicating when the digital assistant has performed a command or is in an inactive state. For example, the electronic device displays a first color for visual indictor **610**A in response to the set of one or more activation conditions being met, displays as second color for visual indicator **610**A in response to the digital assistant being activated, displays a third color for visual indicator **610**A in response to a determination that the digital assistant has performed a command, and/or displays a fourth color for visual indicator **610**A in response to changing the digital assistant to not be activated. An audio indication **610**B optionally includes one or more of: playing a sound, phrase, word, or tone. In some embodiments, audio indicator **610**B indicating that the digital assistant is activated (or that the set of one or more activation criteria has been met) is different from audio indicator **610**B indicating when the digital assistant has performed a command (e.g., **624**A) or is in an inactive state. The digital assistant optionally provides a digital assistant voice confirmation **610**C (e.g., "How can I help you?") indicating that the digital assistant has been activated or that the set of one or more activation criteria has been met. In some examples, as described in further detail with respect to FIGS. **10**A-**10**D, the electronic device outputs the voice confirmation using a particular voice (e.g., accent, male or female voice) based on the detected user.

[0221] In FIG. **6**C, subsequent to the electronic device activating the digital assistant and providing indication(s) using display **602**, the electronic device detects user **620** finishing speaking the remaining portion **624**B " . . . on the table lamp" of command **624**C "Turn on the table lamp,". In some embodiments, the electronic device **600** detects user **620** speaking the entire command **624**C "turn on the table lamp" after the digital assistant is activated. Optionally, electronic device **600** provides indication **610**C in accordance with a determination that the set of one or more activation criteria is met prior to beginning to receive command **624**A-**624**B and electronic device **600** does not provide indication **610**C in accordance with a determination that the set of one or more activation criteria is met after (but not prior) to beginning to receive command **624**A-**624**B.

[0222] In some embodiments, the electronic device **600** is producing an audio output, such as playing music, at a first volume prior to receiving the remaining portion **624**B of command **624**A-**624**B. In some embodiments, in response to electronic device **600** receiving first portion **624**A of command **624**A-**624**B, electronic device **600** reduces (ducks) the first volume of the audio output to a second volume that is lower than the first volume, where the second volume is optionally based on the distance between user **620** and electronic device **600**. The first volume of the audio output is lowered so that electronic device **600** may better detect the words with reduced interference of background noise. In some embodiments, the first volume is reduced in response to the determination that the set of one or more activation criteria is met. In some embodiments, the amount of volume reduction is based on the distance between user **620** and electronic device **600** at the time first portion **624**A of command **624**A-**624**B is received (and not second portion **624**B). Electronic device **600** determines the distance between user **620** and electronic device **600** using the external camera sensors **616**A and/or the camera sensors **616**B on electronic device **600**. In some embodiments, electronic device **600** variably reduces the volume of the audio output by muting or lowering the volume of the audio output while detecting that user **620** is speaking. Electronic device **600** restores the volume of the audio output back to the first volume after detecting the end of command **624**B. The volume change is not based on the content of the request. In some embodiments, even though electronic device **600** detects user **620** speaking the first portion **624**A of command **624**A-**624**B prior to activating the digital assistant, the digital assistant is still able to process the command (including **624**A) if a set of one or more performance criteria is met. When the set of one or more performance criteria is met, the digital assistant performs the command based on the content of the request. The set of one or more performance criteria optionally includes one or more of: a command completion criterion, an actionable criterion, and an activation criteria. The command completion criterion is met when electronic device **600** detects the end of a received audio command (e.g., audio user input request). The end of command **624**A-**24**B is determined by electronic device **600** detecting a pause in received audio (or lack of receiving user audio) for more than a non-zero, predetermined period of time. In some embodiments, the command completion criterion is not met when the detected pause is for less than a non-zero, predetermined period of time (e.g., the user is not done speaking the command). An actionable criterion is met when electronic device **600** determines that the received command is actionable (e.g., that the device has authority and/or capability to perform the command). In some examples, if the command is to access information from an account, but the electronic device has not received authorization to access the account, then the command is not actionable because the device does not have authority to access the account. In some examples, the actionable criterion is not met when electronic device **600** is unable to process the received request into a command that the digital assistant can perform or transmit (e.g., command is to "turn on the tv", but there is no tv in the room to which the corresponding instruction could be sent). The activation criteria is met when the digital assistant has been activated.

[0223] FIG. **6**D illustrates that once the set of one or more performance criteria is met while the digital assistant is activated, the digital assistant performs the command. If the digital assistant successfully processes the command and determines that the performance criteria are met, the digital assistant will send an instruction to cause a change in state of an appropriate device. In this example, the instruction is sent to table lamp **618** (an external device) to turn on. In some embodiments, the determination as to which external device to act on is made in accordance with the description of FIGS. **8**A-**8**C, as described in greater detail below. After the digital assistant has performed the command, electronic device **600** optionally provides an indicator (e.g., indicator **610**A, **610**B, **610**C) indicating whether command **624**A-**624**B has been successfully performed. In some embodiments, the indicator is a visual indicator **610**A that displays a light of a particular color (e.g., blue or a color other than the default color when the digital assistant is not activated) or a light pattern to indicate that the digital assistant has

performed command **624**A-**624**B. In some embodiments, the indicator is an audio indicator **610**B that plays a sound, phrase, word, or tone. In some embodiments, the indicator is a confirmation **610**C spoken by the digital assistant, which can speak in a particular voice (e.g., accent, male or female voice), such as described with respect to FIGS. **10**A-**10**D, when the digital assistant has performed command **624**A-**624**B. In some embodiments, user **620** may customize the visual, audio, or digital assistant indicators to include values of characteristics (e.g., a user name, particular user preferences for light color, pattern, sound, voice, accent) associated with his/her profile as discussed below with respect to FIGS. **10**A-**10**D. For example, in FIG. **6**D, the digital assistant may provide a confirmation, "Jack, the table lamp has been turned on." The name of user **620**, is obtained from the user profile of user **620**, which user **620** provided during registration, and is determined as discussed below with respect to FIGS. **10**A-**10**D. The registration process is described in detail below with respect to FIGS. **10**A-**10**D and with respect to method **1100**. The digital assistant may utilize settings from user **620**'s profile to further customize the indicator to reflect which user (e.g., Jack) has spoken the command

[0224] When the set of one or more performance criteria is not met, electronic device **600** forgoes performing the command based on the content of the command. In some examples, the set of one or more performance criteria is not met when the actionable criterion is not met. When the actionable criterion is not met, electronic device **600** optionally produces a headshake indication as illustrated in FIG. **6**O, to indicate that the digital assistant is unable to perform the command because it could not process the command, lacks authorization, or is incapable of performing the command. The headshake indication in FIG. **6**O comprises a display of lights or a pattern of lights **602**A on a first side of the display **602** of electronic device **600** (as shown in **6**O-A), followed by a pattern of lights **602**B on a second side of the display **602** (as shown in **6**O-B), followed by the pattern of lights **602**A on the first side of the display **602** (as shown in **6**O-C), which simulates a person shaking his/her head. In some embodiments, a particular indicator (e.g., pattern of light, color, or particular sound) is produced by electronic device **600** when the performance criteria is not met. The particular indicator may correspond to a particular error code that indicates why the digital assistant was unable to perform the command (e.g., blue light indicates the command was not understood, purple light indicates lack of authorization). In some embodiments, the indicator produced when the digital assistant is unable to perform the command is different than the indicator produced when the digital assistant is able to perform the command.

[0225] FIGS. **6**E-**6**G illustrate another embodiment where user **620** activates a digital assistant on a wearable or mobile electronic device **630**. In FIG. **6**E, the digital assistant on electronic device **630** is activated when electronic device **630** detects a user input or gesture such as a wrist raise. In some embodiments, user **620** may also optionally press a button, rotate a crown, or provide audio user input to activate the digital assistant on electronic device **630**. Once the digital assistant is activated, electronic device **630** optionally provides a visual indictor **610**A (e.g., light), audio indicator **610**B (e.g., a sound, tone, word), or digital assistant voice confirmation **610**C to indicate that the digital assistant has been activated or that the set of one or more activation criteria are met. The various types of indicators are discussed above with respect to FIG. **6**B.

[0226] In FIG. **6**F, electronic device **630** detects a spoken command comprising command **624**C "Turn on the table lamp" from user **620**. After receiving the spoken command, the digital assistant determines whether a set of one or more performance criteria (e.g., an actionable criterion, a command completion criterion, an activation criterion) are met. As discussed above, when one or more performance criteria is met, the digital assistant performs the command based on the content of the audio input request. As illustrated in FIG. **6**F, the electronic device optionally also determines whether the user's gaze satisfies a set of one or more gaze criteria to identify an external device upon which to act. The set of one or more gaze criteria are described in further detail, below.

[0227] FIG. **6**G illustrates electronic device **630** providing an indicator (e.g., **610**A, **610**B, **610**C) that the command was successfully performed in causing the table lamp to turn on from an off state. The indicators displayed upon the successful performance of the command are discussed above with respect to FIG. **6**E.

[0228] FIGS. **6**H-**6**J illustrate when the digital assistant does not perform the command associated with the spoken command **624**A-**624**B when the set of one or more performance criteria is not met because user **620** has broken his/her gaze while speaking command **642**A-**624**B. The set of one or more performance criteria includes a continuous gaze criterion that is met when electronic device **600** determines that second gaze information (e.g., obtained by electronic device **600** using the one or more camera sensors **616**A, **616**B (while user **620** is speaking) indicates that user **620** did not break his/her gaze directed at electronic device **600** for more than a threshold duration between the activation of the digital assistant and the completion of speaking command **624**A-**624**B. In some embodiments, the threshold duration is zero, meaning the user cannot break his/her gaze while speaking. In some embodiments, the threshold duration is a non-zero time period, which allows the user to briefly look away without cancelling the command.

[0229] In FIG. **6**H-**6**I, user **602** breaks his/her gaze directed at electronic device **600** while speaking command **624**A-**624**B to "turn on the table lamp" while electronic device **600** is activated. As illustrated in FIG. **6**H, the electronic device **600** determines that the user **620** is looking at the electronic device while electronic device **600** receives a first portion **624**A of command **624**A-**624**B "turn . . . ". In FIG. **6**I, electronic device **600** detects a break in the user gaze when user **620** turns 180 degrees to look away from electronic device **600** before having finished speaking the remainder **624**B of command **624**A-**624**B " . . . on the table lamp." The user's gaze is broken from electronic device **600** when electronic device **600** determines that the field of view criterion is not satisfied for more than a non-zero, threshold duration (e.g., 3 seconds).

[0230] FIG. **6**J illustrates that electronic device **600** forgoes performing command **624**A-**624**B to "turn on that light" when the continuous gaze criterion is not met, even when the set of one or more activation criteria has been met (e.g., in FIG. **6**H). As a result, table lamp **618** remains off because the digital assistant has not transmitted the instructions to table lamp **618** to turn on. Electronic device **600** optionally produces an indicator (e.g., **610**A, **610**B, **610**C) that indicates that the command has not been performed. In some embodiments, the indicator may be a visual indicator

610A (e.g., colored light or light pattern), an audio indicator **610**B (e.g., sound, speech, or a particular voice), or a digital voice confirmation **620**C, as discussed above. The indicator (e.g., **610**A, **610**B, **610**C) provided by the electronic device **600** when command **624**A-**624**B is not performed is optionally different than the indicator (e.g., **610**A, **610**B, **610**C) provided by electronic device **600** when command **624**A-**624**B is successfully performed.

[0231] FIGS. **6**K-**6**N illustrate that even if electronic device **600** detects that user **620** breaks his/her gaze directed at electronic device **600** while electronic device receives command **624**D-**624**F (when electronic device **600** is activated), the continuous gaze criterion is still satisfied if the break in the gaze is less than the threshold duration. FIG. **6**K illustrates that electronic device **600** has activated the digital assistant and is receiving portion **624**D "turn . . . " of command **624**D-**624**F to "turn on the table lamp" from user **620**. In FIG. **6**L, electronic device **600** continues to receive second portion **624**E " . . . on the . . . " of command **624**D-**624**F from user **620**, but electronic device **600** detects a break in the user's gaze when the user looks away from electronic device **600**.

[0232] In FIG. **6**M, electronic device **600** detects that user **620** has returned his gaze back to electronic device **600** within the threshold duration (e.g., 3 seconds). The electronic device continues to receive last portion **624**F " . . . table lamp" of command **624**D-**624**F. Because the break in gaze is for less than the threshold duration, the continuous gaze criterion is still satisfied. In FIG. **6**N, when the continuous gaze criterion is satisfied, the digital assistant performs command **624**D-**624**F to "turn on the table lamp" and provides an indication (e.g., visual **610**A, audio **610**B, and digital assistant confirmation **610**C indications) once table lamp **618** has turned on.

[0233] FIGS. **7**A-**7**B are flow diagrams illustrating a method for **700** using an electronic device in accordance with some embodiments. Method **700** is performed at an electronic device (e.g., **100**, **300**, **500**, **600**, **630**, **800**, **1000**). Some operations in method **700** are, optionally, combined, the orders of some operations are, optionally, changed, and some operations are, optionally, omitted.

[0234] As described below, method **700** provides an intuitive way for activating a digital assistant on an electronic device by glancing at the electronic device. The method reduces the cognitive burden on a user when providing user input (e.g., audio input by speaking or touch input via button presses) to activate the digital assistant. In some circumstances, the user is unable to effectively use voice activation (e.g., when the room is noisy) or provide tactile user input (e.g., pressing a button). Therefore, by being able to activate the digital assistant by looking at the electronic device, the user is able to interact with the digital assistant using a more efficient user-device interface.

[0235] While a digital assistant of the electronic device (e.g., **600**) is not activated (**702**), the electronic device (e.g., **600**) obtains (**704**), using one or more camera sensors (e.g., cameras **616**A, **616**B), first gaze information (e.g., of the user **620**). For example, the one or more camera sensors are infrared camera sensor(s) and/or visible light sensor(s). For example, the one or more camera sensors are external to the electronic device (e.g., **600**), such as by being wirelessly connected to the electronic device.

[0236] In accordance with a determination (**706**) (e.g., while the digital assistant is not activated) that the first gaze information satisfies a set of one or more activation criteria: (e.g., the angle of the gaze is less than a threshold number of degrees relative to a line-of-sight, the gaze is in a particular direction, the duration of the gaze is greater than or equal to a predetermined period of time), the electronic device (e.g., **600**) activates (**708**) the digital assistant of the electronic device (e.g., **600**) (e.g., by turning on a microphone, enabling processing of recorded audio).

[0237] Using gaze detection to activate the digital assistant when the set of one or more activation criteria is satisfied provides the user with a more efficient user interface for activating a digital assistant, particularly in contexts where the user cannot press a button or use voice controls because the room is too noisy. Providing improved user interfaces to activate the digital assistant without requiring tactile user input (e.g., a button press) or voice input enhances the operability of the digital assistant and makes the user interface of the digital assistant interface more efficient. In addition, using gaze detection reduces the number of user inputs needed to activate the digital assistant to perform commands, which enhances the operability of the electronic device.

[0238] In some embodiments, the set of one or more activation criteria includes a field of view criterion that is met when a determined offset angle (e.g., **622**) is less than a threshold offset angle (e.g., **604**) (e.g., a non-zero degrees value, the maximum acceptable deviation from the user looking directly at the device), wherein the determined offset angle (e.g., **622**) is an angle formed between: a line-of-sight (e.g., **612**) (e.g., determined from the gaze information) of a gaze of a user (e.g., **620**) and a calculated line-of-sight (e.g., **608**) from the user (e.g., **620**) to the electronic device (e.g., **600**).

[0239] Using a field of view criterion as an activation criteria allows the user's gaze to be detected without requiring the user (e.g., **620**) to look directly at the electronic device (e.g., **600**). Instead, as long as the electronic device (e.g., **600**) is within a field of view (e.g., **606**) of the user's gaze, the user's gaze satisfies the activation criteria, allowing the digital assistant to be activated. Using the field of view criterion allows the digital assistant to be activated via gaze more successfully, which enhances the operability of the digital assistant and makes the gaze detection more efficient (e.g., the digital assistant has a higher activation rate since the electronic device (e.g., **600**) does not have to be in the direct line-of-sight (e.g., **612**) of the user (e.g., **620**)).

[0240] In some examples, if the electronic device (e.g., **600**) is determined to be positioned more than a threshold number of degrees (e.g., outside the field of view **606** of the gaze), the electronic device (e.g., **600**) is not in the user's gaze and the digital assistant is not activated. In some examples, when the electronic device (e.g., **600**) is in the field of view (e.g., **606**) of the gaze by being within the threshold number of degrees, the digital assistant is activated.

[0241] In some embodiments, the line-of-sight (e.g., **612**) of the gaze of the user (e.g., **620**) is based on a head position of the user (e.g., **620**). In some examples, the user's head position may be affected by the tilt or rotation (e.g., yaw, pitch, and roll of the user's head).

[0242] In some embodiments, the set of one or more activation criteria includes a dwell time criterion that is satisfied when a dwell time (e.g., duration) of the gaze determined from the first gaze information is more than a

threshold period of time (e.g., a non-zero threshold period of time). In some examples, if the dwell time of the gaze is less than a threshold time period, the digital assistant is not activated.

[0243] Further in accordance with the determination (**706**) that the first gaze information satisfies the set of one or more activation criteria, the electronic device (e.g., **600**) provides (**710**) an indication that the set of one or more activation criteria has been satisfied and/or that the digital assistant of the electronic device is in transition or transitioning into being activated.

[0244] Providing an indication (e.g., **610**A, **610**B, **610**C) that the set of one or more activation criteria has been satisfied and that the digital assistant is activated provides the user with feedback about the current state of the digital assistant. The indication (e.g., **610**A, **610**B, **610**C) can be visual, audio, or a digital assistant voice confirmation, which provides customizable feedback for various contexts. For example, in contexts where a sound indicator cannot be heard (e.g., a noisy room), providing visual feedback (e.g., **610**A) allows the user (e.g., **620**) to know whether the digital assistant is activated and can process commands. Providing customized feedback of the state of the digital device enhances the operability of the digital assistant and makes the user-device interface more efficient (e.g., the user (e.g., **620**) doesn't issue commands when the digital assistant is not active) by informing the user (e.g., **620**) when the digital assistant is available to process commands (e.g., activated) and when it is unavailable (e.g., inactive). Further, the feedback of the state of the digital assistant allows the user (e.g., **620**) to use the digital assistant more efficiently.

[0245] In some embodiments, the indication (**712**) is a visual indication (e.g., **610**A) (such as the electronic device **600** turning on a display, LED, or light of the electronic device) that indicates that the digital assistant is activated. In some embodiments, the indication (**714**) is an audio indication (e.g., **610**B) (such as electronic device **600** playing a sound through a speaker of the electronic device **600**) to indicate that the digital assistant is now activated.

[0246] In some embodiments, in accordance with a determination (**716**) (e.g., while the digital assistant is not activated) that the first gaze information does not satisfy the set of one or more activation criteria: (e.g., electronic device (e.g., **600**) is not in the user's field of view (e.g., **606**), the gaze is in the wrong direction, the duration of the gaze is less than a predetermined period of time), the electronic device (e.g., **600**) forgoes activating the digital assistant. In some embodiments, forgoing activating the digital assistant includes the digital assistant remaining not activated or the electronic device (e.g., **600**) not enabling the microphone or processing of recorded audio. In some examples, when the first set of gaze information does not satisfy the set of one or more activation criteria, the electronic device (e.g., **600**) also forgoes providing the indicator (e.g., **610**A, **610**B, **610**C) that the set of one or more activation criteria has been satisfied.

[0247] In some embodiments, the electronic device (e.g., **600**) receives (**718**) an audio user input request to perform a command (e.g., **624**A-**624**B, **624**C). In response to the electronic device (e.g., **600**) receiving (**720**) the audio user input request to perform the command (e.g., **624**A-**624**B, **624**C) (or in accordance with the digital assistant interpreting the command (e.g., **624**A-**624**B, **624**C)) and in accordance with a determination that a set of one or more performance criteria (e.g., an actionable criterion, a command completion criterion, an activation criterion) is met, the electronic device (e.g., **600**) performs (**722**) the command (e.g., **624**A-**624**B, **624**C) based on a content of the audio user input request. In some examples, in response to receiving the audio input request and in accordance with a failure to interpret the command (e.g., **624**A-**624**B, **624**C), the electronic device (e.g., **600**) forgoes performing the command (e.g., **624**A-**624**B, **624**C) and provides an indication (e.g., **610**A, **610**B, **610**C) (e.g., a display of a visual indication, such as a headshake indication) that the electronic device (e.g., **600**) did not interpret the command (e.g., **624**A-**624**B, **624**C). Thus, the electronic device (e.g., **600**) provides an indication (e.g., **610**A, **610**B, **610**C) if the digital assistant doesn't understand the audio input. In some examples, when the digital assistant is not activated, the digital assistant forgoes performing the received command (e.g., **624**A-**624**B, **624**C).

[0248] In some embodiments, in response to the electronic device (e.g., **600**) receiving the audio user input request to perform the command (e.g., **624**A-**624**B, **624**C) and in accordance with a determination that the set of one or more performance criteria is not met (e.g., actionable criterion, command completion criterion, activation criterion), the electronic device (e.g., **600**) provides a headshake indication. The headshake indication includes the electronic device (e.g., **600**) displaying a first pattern of lights on a first side of a display (e.g., **602**) of the electronic device (e.g., **600**). Subsequent to displaying the first pattern of lights, the electronic device (e.g., **600**) displays a second pattern of lights on a second side of the display (e.g., **602**) of the electronic device (e.g., **600**). Subsequent to displaying the second pattern of lights, the electronic device (e.g., **600**) displays the first pattern of lights on the first side of the display (e.g., **602**) of the electronic device (e.g., **600**).

[0249] Providing a visual indication (e.g., **610**A) (e.g., a headshake indication (e.g., **602**A-**602**B)) that the set of one or more performance criteria is not met provides the user with feedback that the command issued by the user is unable to be processed. Providing a customizable indication, such as the headshake indication (e.g., **602**A-**602**B), that indicates when the command (e.g., **624**A-**624**B, **624**C) will not be processed enhances the operability of the digital assistant and makes the user-device interface more efficient by informing the user (e.g., **602**) that the command (e.g., **624**A-**624**B, **624**C) will not be performed. In addition, customized indications (e.g., **610**A, **610**B, **610**C) corresponding to error codes provide improved feedback to the user which communicates why the command (e.g., **624**A-**624**B, **624**C) was not carried out, which makes the user-device interface more efficient.

[0250] In some embodiments, the set of one or more performance criteria includes a completion criterion that is met when the electronic device (e.g., **600**) detects the end of a received audio command (e.g., the end of the audio user input request, a pause of audio for more than a predetermined period of time). In some examples, the set of one or more performance criteria includes a actionable criterion that is met when the electronic device (e.g., **600**) determines that the received command is actionable (e.g., that the digital assistant has authority and/or capability to perform the command).

[0251] In some embodiments, the electronic device (e.g., **600**) starts receiving the audio user input request while the

digital assistant is not activated. In some examples, the digital assistant is inactive because the gaze information does not yet satisfy the set of one or more activation conditions. This may happen, for example, when the user (e.g., **620**) begins speaking the command (e.g., **624**A-**624**B, **624**C) before looking at the digital assistant. In some examples, first gaze information satisfies the set of one or more activation criteria and, in response, the digital assistant is activated before the electronic device (e.g., **600**) finishes receiving the audio user input request to perform the command.

[0252] Being able to receiving the audio user input request before the digital assistant is activated allows the user (e.g., **602**) to begin speaking even if the user (e.g., **602**) hasn't activated the digital assistant yet. Detecting the audio user input request when the set of activation conditions has not yet been met enhances the operability of the digital assistant and makes the device-user interface more efficient because the activation and speaking of a command (e.g., **624**A-**624**B, **624**C) do not have to be performed in order.

[0253] In some embodiments, in response to the electronic device (e.g., **600**) receiving the audio user input request to perform the command (e.g., **124**C) and in accordance with a determination that the set of one or more performance criteria is not met (e.g., various criteria that should be met for the electronic device (e.g., **600**) to perform the command (e.g., **624**A-**624**B, **624**C)), the electronic device (e.g., **600**) forgoes (**724**) performing the command (e.g., **624**A-**624**B, **624**C) based on the content of the audio user input request, wherein the set of one or more performance criteria includes a continuous gaze criterion that is met when the electronic device (e.g., **600**) determines that second gaze information (e.g., obtained by the electronic device using the one or more camera sensors) indicates that the user (e.g., **602**) does not break a gaze directed at the electronic device (e.g., **600**) for more than a threshold duration (e.g., a non-zero duration) between activation of the digital assistant and the completion of the audio user input request. In some examples, the digital assistant does not perform the command (e.g., **624**A-**624**B, **624**C) because the user (e.g., **602**) breaks their gaze with the device (e.g., user (e.g., **602**) looked away before completion of the command, user (e.g., **602**) looked away for more than 1 second at a time). In some examples, the threshold duration is 0 seconds. In some examples, the threshold duration is a duration that is non-zero (e.g., 0.5 seconds).

[0254] In some embodiments, the electronic device (e.g., **600**) produces an audio output (e.g., the electronic device (e.g., **600**) is playing a song), wherein the audio output is being produced at a first volume prior to receiving the audio user input request. In response to beginning to receive the audio user input request, the electronic device (e.g., **600**) reduces a volume of the audio output to a second volume that is lower than the first volume, wherein the second volume is based on a distance between a user (e.g., the user providing the audio user input request) and the electronic device (e.g., **600**). In some examples, the electronic device (e.g., **600**) determines a distance between a user (e.g., **602**) and the electronic device (e.g., **600**) using the external sensors (e.g., **616**A) or the sensor (e.g., **616**B) on the electronic device (e.g., **600**) and determines whether to mute or reduce the volume of the music while the user (e.g., **602**) is talking (e.g., variably duck the audio output). In some examples, the electronic device may bring the volume back up to the first volume after detecting the end of the command (e.g., **624**C). In some examples, the volume change is not based on the content of the request.

[0255] Reducing the audio output volume or muting the audio output when the electronic device receives an audio user input request enables the audio user input request to be more easily processed and results in a higher likelihood that the command (e.g., **624**C) derived from the audio user input request will be understood by the electronic device as ambient noise is reduced. Removing possible ambient noise or sounds that could interfere with the command (e.g., **624**C) that is processed by the digital assistant enhances the operability of the digital assistant and makes the user-device interface more efficient by enabling a higher accuracy of executing the correct command, which improves the user experience as the user (e.g., **602**) doesn't have to repeat the command again when the digital assistant is unable to process the command (e.g., **624**C).

[0256] Note that details of the processes described above with respect to method **700** (e.g., FIGS. **7**A-**7**B) are also applicable in an analogous manner to the methods described below. For example, method **700** may include one or more of the characteristics of the various methods described below with reference to the processes in FIGS. **9**A, **9**B, **11**A, and **11**B. For brevity, these details are not repeated below.

[0257] It should be understood that the particular order in which the operations in the following FIGS. **9**A-**9**B and **11**A-**11**B have been described is exemplary and not intended to indicate that the described order is the only order in which the operations could be performed. One of ordinary skill in the art would recognize various ways to reorder the operations described herein, as well as excluding certain operations. For brevity, these details are not repeated here. Additionally, it should be noted that aspects of the methods and processes described throughout this description may be incorporated with one another.

[0258] FIGS. **8**A-**8**L illustrate exemplary user interfaces for the electronic device to use gaze information to determine the context (e.g., an appliance that the command is directed at) for performing a command once the digital assistant has been activated, in accordance with some embodiments. The techniques illustrated in FIGS. **8**A-**8**L optionally work in conjunction with or include the techniques illustrated in FIGS. **6**A-**6**O, which disclose how the user may activate the digital assistant by looking at the electronic device and issuing commands. Thus, the techniques described below include one or more of the characteristics of the various techniques described above with reference to FIGS. **6**A-**6**O and method **700**. The user interfaces in these figures are used to illustrate the processes described below, including the processes in FIGS. **9**A-**9**B.

[0259] FIGS. **8**A-**8**C illustrate electronic device **800** receiving an audio user input request from user **820** to perform a first command **824**A-**824**B once the digital assistant of electronic device **800** has been activated. The digital assistant is activated based on a trigger word, a gaze (e.g., as discussed above with respect to FIGS. **6**A-**6**O), a button press, a wrist raise (FIG. **6**E-**6**G) while the user is wearing a wearable electronic device (e.g., a watch), and/or detecting an external device (e.g., a phone) being pointed at the electronic device **800**. The electronic device (e.g., **800**) optionally provides a visual or audio indication **810**A-**810**B to indicate that the digital assistant is activated.

[0260] In FIG. **8**A, user **820** is standing in a room with a table lamp **818** and a floor lamp **814**. The digital assistant of electronic device **800** has been activated based on the electronic device **800** determining that the set of one or more activation criteria have been satisfied with respect to method **700** as discussed above (e.g., device **800** determines that user **820** is looking at device **800**). Once the digital assistant is activated, the electronic device detects user **820** speaking a portion **824**A "turn . . . " of the full command **824**A-**824**B. In some embodiments, the complete command **824**A-**824**B is applicable to the multiple external devices (e.g., table lamp **818**, floor lamp **814**) that may be in the room. In some embodiments, the user's command **824**A-**824**B does not include enough detail for the digital assistant to determine which external device (e.g., table lamp **818**, floor lamp **814**) the command **824**A-**824**B is directed to. Therefore, based on the command **824**A-**824**B alone, the electronic device **800** is unable to determine whether "that light" refers to table lamp **818** or floor lap **814**.

[0261] In FIG. **8**B, user **820** indicates which external device (e.g., light) he/she wants to act on (e.g., turn on) by looking at table lamp **818**, while finishing the remaining portion **824**B " . . . on that light" of command **824**A-**824**B "turn on that light." The electronic device receives and uses gaze information based on the user's gaze, while the user **820** is speaking, to determine that table lamp **818** is the intended device, when the user's gaze satisfies a set of one or more gaze criteria. Electronic device **800** obtains information relating to the user's gaze through the one or more camera sensors that may be located on external devices **816**A and/or is integrated with electronic device **800** at camera **816**B. Cameras **816**A are wirelessly connected to electronic device **600**, though they could alternatively be wired. The one or more camera sensors include infrared camera sensors and/or visible light sensors and measure gaze information once a gaze has been detected. The gaze information optionally includes one or more of: the position of the user's head position, the dwell time or duration of the gaze, the direction of the gaze, the field of view **806** of the gaze, the direction of the gaze, and whether an object (e.g., electronic device **800** or an external device (e.g., lamps **814**, **818**) is within the field of view **806** of the gaze (e.g., external device is within a threshold offset angle **804** from a user's line-of-sight **812**).

[0262] In some embodiments, electronic device **800** determines that the set of one or more gaze criteria is met when a direction criterion, a dwell time criterion, and/or a field of view criterion is met. The set of one or more gaze criteria optionally includes a direction criterion that is satisfied when the gaze is directed in the direction of table lamp **818**.

[0263] The set of one or more gaze criteria optionally includes a dwell time criterion that is satisfied when electronic device **800** detects a dwell time or duration of the gaze in the direction of table lamp **818** for greater than a non-zero, predetermined period of time (e.g., two seconds, a threshold time period). In some examples, electronic device **800** utilizes the dwell time to determine whether user **820** is looking at table lamp **818** in order to determine which external object to act upon (e.g., to send the instruction to). If the dwell time or duration of the gaze is too short or less than the threshold time period, this indicates that the user does not intend the particular external device (e.g., table lamp **818**) to be the target device of the command **824**A-**824**B. In some embodiments, a dwell time less than the threshold time period does not satisfy the gaze criteria, resulting in the command **824**A-**824**B not being transmitted to the particular external device identified by electronic device **800**. In some embodiments, a dwell time less than the threshold time period results in electronic device **800** not identifying any external device.

[0264] The set of one or more gaze criteria optionally includes a field of view criterion. The field of view criterion is satisfied for a particular external device when an external device, such as table lamp **818**, is determined to be in field of view **806** of user **820**. Table lamp **818** is determined to be in field of view **806** of user **820** when table lamp **818** is within a threshold offset angle **804** from a user's line-of-sight **812**. Using the field of view to determine whether a user is looking at a target external device (e.g., table lamp **818**, floor lamp **814**) results in a higher gaze detection rate as a gaze may be detected in the field of view even if the user is not looking directly at the target device. In some embodiments, the field of view criterion is satisfied for table lamp **818** when table lamp **818** is directly in the line-of-sight **812** of user **820**. In some embodiments, the field of view criterion is satisfied for a particular external device when a determined offset angle of the gaze is less than the threshold offset angle **804** to the external device (e.g., the maximum acceptable deviation from a line-of-sight **812** of user **820** looking directly at table lamp **818**). The threshold offset angle **804** is greater than zero degrees (e.g., thirty degrees). The determined offset angle is calculated based on the angle formed between a line-of-sight **812** from a gaze of user **820** and a calculated line-of-sight from user **820** to electronic device **800**. In the example of FIG. **8**B, user **820** is looking directly at table lamp **818**. In some examples, when table lamp **818** is in the field of view of the gaze by being within the threshold number of degrees from the user's direct line-of-sight, the field of view criterion is met for table lamp **818** and electronic device **800** optionally transmits the instructions to table lamp **818**. When the offset angle is greater than threshold offset angle **804**, the field of view criterion is not satisfied resulting in the digital assistant not transmitting the instructions to table lamp **818**, which is outside the field of view **806** of user **820**. In some embodiments, if table lamp **818** is determined to be positioned more than a threshold number of degrees outside the field of view **808** of user **820**, table lamp **818** is determined to not be in the user's gaze and electronic device **800** does not transmit the instructions to table lamp **818**. In some embodiments, line-of-sight **812** is measured based on the head position of the user. In some embodiments, the head position of the user is based on the tilt or rotation (e.g., yaw, pitch, and/or roll) of the user's head. Aspects of these techniques are illustrated and described in further detail above with respect to FIGS. **6**B-**6**C.

[0265] Electronic device **800** optionally provides an indication (e.g., indicators **810**A, **810**B, **810**C) when the set of one or more gaze criteria is met for table lamp **818**. The indication optionally indicates that the command **824**A-**824**B from the received audio user input request is associated with or is directed to the table lamp **818**. The indication may be a visual **810**A or audio **810**B indicator that indicates that the digital assistant has identified table lamp **818** based on the user's gaze. In some embodiments, the indicator is a visual indicator **810**A that displays a light of a particular color (e.g., blue or a color other than the default color when table lamp **818** is identified) or a light pattern to indicate that

the digital assistant has identified table lamp **818** based on the user's gaze. In some embodiments, the indicator is an audio indicator **810**B that plays a sound, word, tone, or speaks in a particular voice (e.g., accent, male or female voice) when the digital assistant has identified table lamp **818**. In some embodiments, the digital assistant provides an audio confirmation ("turning on the table lamp") to indicate that electronic device **800** has identified the external device to act on (e.g., to turn on).

[0266] In FIG. **8**C, when the electronic device determines that the set of one or more gaze criteria is met for table lamp **818**, the digital assistant transmits instructions to act on the command. For example, the electronic device determines that the command is to perform a function "turn on" and that the function should be directed at a particular device—the table lamp **818**. Accordingly, the electronic device transmits an instruction (e.g., to table lamp **818** or to another device that controls table lamp **818**) to turn the table light from an off state to an on state. In some embodiments, the digital assistant transmits instructions that causes an external device to change tasks as a result of transitioning from a first state to a second state. For example, the digital assistant may transmit instructions that causes an external device (e.g., a computer or phone) to transition from a first state to a second state when the external device changes a song, transitions from performing one task to another (e.g., playing music to providing the weather), or transitions from an inactive state to an active state, and vice versa.

[0267] As illustrated in FIG. **8**C, when there are multiple external devices (e.g., table lamp **818**, floor lamp **814**) in the room (or accessible by electronic device **800**), electronic device **800** is able to differentiate between the multiple external devices and only transmits instructions to the intended external device (e.g., table lamp **818**, floor lamp **814**). Electronic device **800** identifies the intended external device (e.g., table lamp **818**, floor lamp **814**) by determining whether the set of one or more gaze criteria is met for a particular external device (e.g., table lamp **818**, floor lamp **814**). In other words, electronic device **800** uses the gaze information to determine which external device (e.g., table lamp **818**, floor lamp **814**) the user is looking at, when the set of one or more gaze criteria is satisfied. In some embodiments, when electronic device **800** transmits the instructions (e.g. "turn on") to the identified first external device, table lamp **818**, a second external device, floor lamp **814** in the room remains unaffected because floor lamp **814** is not the external device identified by the gaze information. Floor lamp **814** (the second external device) may be in an on state or an off state. For example, in FIG. **8**C, the instructions are transmitted to table lamp **818** causing table lamp **818** to turn on, but the instructions are not transmitted to floor lamp **814**, which remains off (floor lamp **814** was previously in the off state). Thus the transmitted instruction does not cause floor lamp **814** to change state. In some embodiments, when user **820** glances at floor lamp **814** instead of table lamp **818**, the set of one or more gaze criteria is met by floor lamp **814**, instead of table lamp **818**. Accordingly, the digital assistant transmits the instructions to floor lamp **814**, while table lamp **818** remains unaffected (e.g., table lamp **818** remains on if it was previously on, and remains off if it was previously off).

[0268] When the digital assistant has transmitted the instructions to the corresponding external device, electronic device **800** optionally provides an indicator (e.g., indicator **810**A, **810**B, **810**C) indicating that the table lamp **818**, has transitioned from a first state to a second state. In some embodiments, the indicator is a visual indicator **810**A that displays a light of a particular color (e.g., blue or a color other than the default color) or a light pattern to indicate that table lamp **818** has turned on, from the off state, as a result of receiving the instructions. In some embodiments, the indicator is an audio indicator **810**B that is a sound, one or more words, or a tone. In some embodiments, the indicator is an audio confirmation **810**C from the digital assistant, such as "Ok, Jack, the table lamp is on." The user may customize the indicators to include characteristics (e.g., a user name, particular user preferences for light color, pattern, sound, voice, accent) associated with the user's profile, which user **820** provided during registration. The registration process is described in detail below with respect to FIGS. **10**A-**10**D and method **1100**. The digital assistant customizes the indicator to reflect which user (e.g., Jack) was recognized, as discussed below with respect to FIGS. **10**A-**10**D and method **1100**. In some embodiments, electronic device **800** provides the indicator (e.g., **810**A, **810**B, **810**C) when the digital assistant has successfully sent instructions to table lamp **818**, regardless of whether table lamp **818** actually turned on. In some embodiments, electronic device **800** provides the indicator (e.g., **810**A, **810**B, **810**C) in response to a determination that table lamp **818** has turned on. Optionally, electronic device **800** provides a different indicator (e.g., **810**A, **810**B, **810**C) when table lamp **818** was unable to execute the command to "turn on" and remains off. In some embodiments, the electronic device **800** may provide a different indicator (e.g., **810**A, **810**B, **810**C) to indicate a particular error code when the command **824**A-**824**B was unsuccessfully executed.

[0269] The indicators (e.g., **810**A, **810**B, **810**C) are optionally domain (e.g., music) specific. In some embodiments, the electronic device **800** determines that the command refers to an action in a particular domain (e.g., "Love that song" command in the music domain). In accordance with a determination that the command is of a first domain (e.g., music), the indication is a first type of indication (e.g., a particular set of sounds or light indictors). In accordance with a determination that the command is of a second domain (e.g., e-mail) different from the first domain, the indication is a second type of indication (e.g., a different set of sounds or light indicators) different from the first type of indication. Having different indicators (e.g., **810**A, **810**B, **810**C) for different domains allows the user to easily recognize when a command was executed successfully versus when a command failed to execute.

[0270] FIG. **8**D illustrates that once the digital assistant has transmitted the instruction to turn on table lamp **818**, user **820** may further change a setting of the external device (e.g., the brightness of the lamp) by providing an input at a second external device (e.g., by rotating the crown **832** of a watch **830** that user **820** is wearing). Watch **830** does not have access to contextual information about which external device to control. Instead, the user merely uses the crown **832** of watch **830** as an extended input mechanism to control the lamp. Watch **830** detects the user input and sends an indication of the user input to electronic device **800**. Electronic device **800** receives the indication of the user input and determines that table lamp **818** is the intended external device (e.g., table lamp **818**, floor lamp **814**) that should receive the instructions. In FIG. **8**D, user **820** turns the

crown **842** to increase the brightness of the table lamp **818** to a brightness of 80. In some examples, electronic device **800** determines that table lamp **818** is the intended external device based on the recency of electronic device **800** having transmitted instructions to table lamp **818** based on user **820**'s actions. In some embodiments, the electronic device **800** transmits the instructions (based on the user input at watch **830**) to the most recent device (e.g., table lamp **818**) that the electronic device **800** instructed. For example, if the electronic device **800** had just instructed the computer to play music, then the user turning the crown **842** on the watch **840**, results in electronic device **800** transmitting instructions to the computer to increase (or decrease) the volume of the computer. In some embodiments, electronic device **800** is paired to table lamp **818**. In some embodiments, the electronic device **800** detects that user **820** is wearing watch **840** while watch **840** is paired with the electronic device **800**. In some embodiments, watch **840** is (also) paired with the table lamp **818**. Once electronic device **800** has sent instructions to table lamp **818**, electronic device **800** provides an indicator (e.g., **810**A, **810**B, **810**C) to indicate that the instructions were successfully sent. In some embodiments, electronic device **800** provides a different indicator (e.g., **810**A, **810**B, **810**C) indicating that the instructions were not successfully sent to table lamp **818**.

[0271] FIG. **8**E illustrates the electronic device **800** receiving an additional indication of user input from watch **830** to dim the light of table lamp **818** to a lower brightness, which corresponds to an amount of rotation of crown **832** by user **820**. In response to receiving the indication of the input from watch **830**, electronic device **800** transmits an instruction to table lamp **818** to dim the lights from a brightness of 80 to a reduced brightness of 20, based on the recency of having instructed table lamp **818** to turn on. In some embodiments, user **820** uses watch **830** to directly control table lamp **818** once table lamp **818** has been turned on without having to provide additional commands to the digital assistant. Once electronic device **800** has sent instructions to table lamp **818**, electronic device **800** provides an indicator (e.g., **810**A, **810**B, **810**C) to indicate that the instructions were successfully sent. In some embodiments, electronic device **800** provides a different indicator (e.g., **810**A, **810**B, **810**C) indicating that the instructions were not successfully sent to table lamp **818**.

[0272] In some embodiments, as illustrated in FIGS. **8**F-**8**H, when there are multiple external devices (e.g., floor lamp **814**, table lamp **818**) in the room, electronic device **800** uses the user's gaze information to identify a particular external device as the intended external device based on the user's gaze. Electronic device **800** identifies the intended external device by determining whether the set of one or more gaze criteria is met for the particular external device based on the gaze information. When the set of one or more gaze criteria is not satisfied for any external device, electronic device **800** forgoes transmitting the instruction because electronic device **800** is unable to determine the external device from the gaze information. As a result, the external devices in the room remain in their original states.

[0273] FIG. **8**F illustrates electronic device **800** detecting gaze information of user **820** and, in accordance with a determination that the set of one or more activation criteria has been satisfied (e.g., as discussed above with respect to FIGS. **6**A-**6**D), the electronic device activates the digital assistant. The electronic device may optionally provide a visual indicator **810**A (e.g., turning on display **802**), audio indicator **810**B, or a digital voice confirmation **810**C to indicate that the digital assistant is activated.

[0274] FIG. **8**G illustrates a table lamp **818** that is initially on and floor lamp **814** that is initially off. Electronic device **800** detects user **820** speaking a second command **824**C to "turn on that light." However, user **820** does not look at floor lamp **814**, which is the intended external device of command **826**C. Electronic device **800** obtains second gaze information through the one or more camera sensors **816**A that may be located on external devices and/or is integrated with electronic device **800** at camera **816**B. The second gaze information optionally includes one or more of: the position of the user's head position, the dwell time or duration of the gaze, the direction of the gaze, the field of view **806** of the gaze, the direction of the gaze, and whether floor lamp **814** is within the field of view **806** of the gaze.

[0275] In FIG. **8**H, electronic device **800** determines that the second gaze information obtained from the one or more camera sensors **816** does not meet the set of one or more gaze criteria. As discussed above, the set of one or more gaze criteria optionally includes a direction criterion that is satisfied for floor lamp **814** when the gaze is directed in the direction of floor lamp **814**. The set of one or more gaze criteria optionally includes a dwell time criterion that is satisfied for floor lamp **814** when electronic device **800** detects a dwell time or duration of the gaze in the direction of floor lamp **814** for greater than a predetermined period of time (e.g., a non-zero threshold, two seconds). The set of one or more gaze criteria optionally includes a field of view criterion. The field of view criterion is satisfied for floor lamp **814** when floor lamp **814** is determined to be in field of view **808** of user **820**. FIGS. **8**G-**8**H illustrates that floor lamp **814** and table lamp **818** are not determined to be in field of view **808** of user **820** because the determined offset angles is greater than the threshold offset angle from user **820**'s line-of-sight **812**. The threshold offset angle **804** is greater than zero degrees (e.g., thirty degrees). When the set of one or more gaze criteria is not met because one or more of a duration criterion, dwell time criterion, and/or field of view criterion are not met for any external device, electronic device **800** is unable to determine which external device (floor lamp **814**, table lamp **818**) user **820** intends to act on (e.g., to turned on).

[0276] FIG. **8**H illustrates that because the set of one or more gaze criteria was not met for any external device, electronic device **800** forgoes transmitting the instruction "turn on that light" to transition the external device, floor lamp **814** or table lamp **818**, from the first state (e.g., off) to the second state (e.g., on). As a result, floor lamp **814** remains off and table lamp **818** remains off, which was the initial states of the external devices prior to electronic device **800** detecting user **820** speaking the command **824**C. For example, table lamp **818** remains in its on state because no instructions have been sent to table lamp **818** to change its initially on state.

[0277] Electronic device **800** optionally produces a particular indicator (e.g., **810**A, **810**B, **810**C) (e.g., a sequence of lights such as a headshake indication (e.g., FIG. **6**O or sounds) to indicate that electronic device **800** did not transmit the instructions because it could not identify the floor lamp **814** (or any other device) as the intended external device based on the gaze information of user **820**. In some embodiments, a particular indicator (e.g., pattern of light,

color, or particular sound) may correspond to a particular error code for why the digital assistant was unable to determine the external device associated with command **824**C. The indicator produced when the digital assistant is unable to transmit the instructions is optionally different than the indicator produced when the digital assistant is able to transmit the instruction. In some embodiments, no indicator is produced when the digital assistant does not transmit the instructions associated with command **824**C.

[0278] FIG. **8**I illustrates another embodiment where there are multiple external devices (e.g., floor lamp **814** and table lamp **818**) in the room and paired to electronic device **800**. In this example, the digital assistant has been activated by user **820**'s glance. In other examples, the digital assistant is activated by a word, a glance, or by a button press, as discussed above. As shown in FIG. **8**I, prior to the user speaking command **824**A-**824**B in FIGS. **8**J-**8**L, table lamp **818** is already in an on state, while floor lamp **814** is in an off state.

[0279] In FIG. **8**J, electronic device **800** detects user **820** speak a portion **824**A, "Turn . . . ", of a second command **824**A-**824**B to "turn on that light" while detecting that user **820** is looking at floor lamp **814**. At FIG. **8**J, electronic device **800** obtains second gaze information based on the user's glance at floor lamp **814** through the one or more camera sensors **816**A that may be located on external devices and/or is integrated with electronic device **800** at camera **816**B. The second gaze information optionally includes one or more of: the position of the user's head position, the dwell time or duration of the gaze, the direction of the gaze, the field of view of the gaze, the direction of the gaze, and whether floor lamp **814** is within the field of view of the gaze based on an offset angle.

[0280] Electronic device **800** determines whether the second gaze information obtained from the one or more camera sensors **816**A-**816**B satisfies a set of one or more gaze criteria. When the set of one or more gaze criteria is met for floor lamp **814**, electronic device **800** is able to determine that user **820** intended the second command **824**A-**824**B to be applied to floor lamp **814**. As discussed above, the set of one or more gaze criteria optionally includes a direction criterion that is satisfied for floor lamp **814** when the gaze is directed in the direction of floor lamp **814**. The set of one or more gaze criteria optionally includes a dwell time criterion that is satisfied when electronic device **800** detects a dwell time or duration of the gaze in the direction of floor lamp **814** for greater than a non-zero, predetermined period of time (e.g., two seconds). The set of one or more gaze criteria optionally includes a field of view criteria. The field of view criterion is satisfied when external device, floor lamp **814**, is determined to be in field of view **808** of user **820**. Floor lamp **814** is determined to be in field of view **808** of user **820** when floor lamp **814** is within a threshold offset angle **804** from a user's line of-sight **812**. In some embodiments, the field of view criterion is satisfied when floor lamp **814** is directly in the line-of-sight **812** of user **820**, as shown in FIG. **8**J. In some embodiments, a field of view criterion is satisfied when a determined offset angle of the gaze is less than the threshold offset angle (e.g., the maximum acceptable deviation from a line-of-sight **812** of user **820** looking directly at floor lamp **814**). The threshold offset angle **804** is optionally greater than zero degrees (e.g., thirty degrees). The determined offset angle is calculated based on the angle formed between a line-of-sight **812** from a gaze of user **820** and a calculated line-of-sight **808** from user **820** to electronic device **800**. In some examples, when floor lamp **814** is in the field of view of the gaze by being within the threshold number of degrees from the user's direct line-of-sight **812**, the field of view criterion is met for floor lamp **814**.

[0281] In some embodiments, as shown in FIG. **8**K, user **820** breaks his gaze at the floor lamp **814** while speaking the remaining portion **824**B of the second command **824**A-**824**B. FIG. **8**K illustrates electronic device **800** detecting that user **820** is looking away from floor lamp **814**, thereby breaking his gaze with floor lamp **814** while speaking the second portion **824**B of second command **824**A-**824**B. In some embodiments, when electronic device **800** has already determined that floor lamp **814** is the intended external device prior to user **820** looking away, the set of one or more gaze criteria is still met even though user **820** has broken his/her gaze prior to completing second command **824**A-**824**B. In some embodiments, if electronic device **800** detects that user **820** returns his gaze back to floor lamp **814** within a threshold duration, prior to completing command **824**A-**824**B (e.g., FIGS. **6**E-**6**G) electronic device **800** still identifies the floor lamp **814** as the external device.

[0282] When the set of one or more gaze criteria is met, the electronic device transmits the instruction to floor lamp **814**, causing floor lamp **814** to transition from an off state to an on state, as shown in FIG. **8**L. Table lamp **818** remains in the prior on state, because table lamp **818** does not meet the set of one or more gaze criteria, and therefore, the electronic device does not transmit instructions for changing the state of the table lamp **818**. In some embodiments, if the command **824**C received by electronic device **800** was determined to be directed to floor lamp **814** and included "turn off that light," when the set of one or more gaze criteria is met for floor lamp **814**, electronic device **800** would transmit instructions to turn off floor lamp **814**, while table lamp **818** that was previously in the on state would remain on.

[0283] FIG. **8**L illustrates when electronic device **800** has transmitted the instruction to floor lamp **814**, electronic device **800** optionally provides an indicator (e.g., indicator **810**A, **810**B, **810**C) indicating that floor lamp **814** has transitioned from an off state to an on state. In some embodiments, the indicator is a visual indicator **810**A that displays a light of a particular color (e.g., blue or a color other than the default color) or a light pattern to indicate that table lamp **818** has turned on, from the off state, as a result of receiving the instructions. In some embodiments, the indicator is an audio indicator **810**B that plays a sound, one or more words, or a tone. In some embodiments, the digital assistant speaks a confirmation **810**C. In some embodiments, user **820** may customize the audio and visual indicators to include characteristics (e.g., a user name, particular user preferences for light color, pattern, sound, voice, accent) associated with his/her user profile. In FIG. **8**K, the digital assistant provides an audio indicator, "Jack, floor lamp turned on." The audio indicator provides a confirmation that the instruction was executed. The audio indicator also includes the name of user **820**, which was obtained from the user profile of user **820**, which user **820** provided during registration. The registration process is described in detail below with respect to FIGS. **10**A-**10**D. The digital assistant may utilize settings from user **820**'s profile to further customize the indicator to reflect which user (e.g., Jack) has spoken the command. In some embodiments, electronic device **800** provides the indicator when the electronic device

has successfully sent instructions to "turn on" to floor lamp **814**, regardless of whether floor lamp **814** actually turned on. Optionally, electronic device **800** provides a different indicator when floor lamp **814** was unable to execute the command to "turn on" and remains off. In some embodiments, the indictor when the command **824**C was unsuccessfully executed includes different indicators for different error codes.

[0284] FIGS. **9**A-**9**B are flow diagrams illustrating a method for **900** using an electronic device in accordance with some embodiments. Method **900** is performed at an electronic device (e.g., **100**, **300**, **500**, **600**, **640**, **800**, **1000**). Some operations in method **900** are, optionally, combined, the orders of some operations are, optionally, changed, and some operations are, optionally, omitted.

[0285] As described below, method **900** provides an intuitive way for the user (e.g., **820**) to provide context for an audio user input request to a digital assistant on an electronic device (e.g., **800**) by glancing at the external device (e.g., **814**, **818**) that the audio user input request is intended to operate on. The method reduces the cognitive burden on a user (e.g., **820**) when providing audio user input request (e.g., having to remember to say the external device in the command) to the digital assistant. Therefore, the digital assistant is able to determine the external device (e.g., **814**, **818**) for the command by monitoring the gaze of the user (e.g., **820**).

[0286] While a first external device (e.g., **818**) is in a first state (**902**) (e.g., the first external device (e.g., **818**) is on or off, the first external device (e.g., **818**) is playing a particular media item) the electronic device (e.g., **800**) receives (**904**) an audio user input request (e.g., a user speech input) to perform a first command (e.g., **824**A-**824**B) (e.g., to transition the external device to a second state, telling the digital assistant to "turn on that light"). In some embodiments, the electronic device receive (**904**) the audio user input request while the digital assistant on the electronic device (e.g., **800**) is active (e.g., based on a trigger word, gaze, and/or button press).

[0287] While the first external device (e.g., **818**) is in the first state (**902**), the electronic device obtains (**906**), using one or more camera sensors (e.g., **816**A, **816**B) (e.g., infrared camera sensor(s), visible light sensor(s)); from camera (s) external to the electronic device (e.g., **800**) such as wirelessly connected to the electronic device or part of the electronic device (e.g., **800**)), first gaze information (e.g., the electronic device (e.g., **800**) detects gaze information that is directed at the first external device (e.g., **818**)). In some examples, the electronic device (e.g., **800**) determines, using the first gaze information, whether a set of one or more gaze criteria is met for a particular external device.

[0288] In accordance with a determination (**908**) (and, optionally, in response to the electronic device (e.g., **800**) receiving the audio user input request), using the first gaze information, that a set of one or more gaze criteria is met for the first external device (e.g., **818**) (e.g., appliance such as a lamp, tv, or computer): the electronic device (e.g., **800**) transmits (**910**) (e.g., to the first external device (e.g., **818**)), based on the first command, an instruction to transition the first external device (e.g., **818**) from the first state to a second state (e.g., the digital assistant performs an action on the external device (e.g., **818**) by causing the external device (e.g., **818**) to transition from an ON state to an OFF state or from an OFF state to an ON state).

[0289] Using gaze detection to determine which external device (e.g., **814**, **818**) is associated with the command (e.g., **824**A-**824**B) when the set of one or more gaze criteria is satisfied provides the user with a more efficient user interface for providing context to the digital assistant simply by looking at the external device (e.g., **814**, **818**) without having to answer follow up questions from the digital assistant to clarify which external device (e.g., **814**, **818**) is intended. Providing improved user interfaces to provide context to a command (e.g., **824**A-**824**B) to the digital assistant without the digital assistant having to interact with the user to obtain more information makes the user interface of the digital assistant interface more efficient. In addition, using gaze detection reduces the number of inputs needed to activate the digital assistant to perform commands (e.g., **824**A-**824**B), which enhances the operability of the digital assistant.

[0290] In some embodiments, in accordance with the determination that the set of one or more gaze criteria is met for the first external device (e.g., **818**) (e.g., angle, duration, and direction of gaze), the electronic device (e.g., **800**) provides (**912**) an indication (e.g., **810**A, **810**B, **810***c*) (e.g., visual or audio indication) that the set of one or more gaze criteria is met (e.g., indicating that the first command (e.g., **824**A-**824**B) is associated with the first external device (e.g., **818**)).

[0291] Providing an indication that the set of one or more gaze criteria has been satisfied and that the digital assistant has successfully transmitted the instructions to the external device (e.g., **800**) provides the user with feedback about whether the command (e.g., **824**A-**824**B) was executed. The indication (e.g., **810**A, **810**B, **810**C) can be visual, audio, or a digital assistant voice confirmation, which provides customizable feedback for various contexts. For example, in contexts where sound cannot be heard (e.g., a noisy room), providing visual feedback allows the user to know whether the digital assistant has transmitted the command (e.g., **824**C). Providing customized feedback of the state of the digital device enhances the operability of the digital assistant and makes the user-device interface more efficient (e.g., the user doesn't have to repeat the command because the user is not sure if it was carried out by the digital assistant) by informing the user when the digital assistant has successfully transmitted the command to the external device. Further, the feedback of the state of the digital assistant allows the user to use the device more quickly and efficiently.

[0292] In some embodiments, the indication (e.g., **810**A, **810**B, **810**C) (e.g., visual or audio indication) is domain specific (**914**) (e.g., music domain, weather domain, messaging domain). In some examples, the electronic device (e.g., **800**) determines that the command (e.g., **824**C) is of a particular domain. In accordance with a determination that the command (e.g., **824**C) is of a first domain, the indication (e.g., **810**A, **810**B, **810**C) is a first type of indication. In accordance with a determination that the command (e.g., **824**C) is of a second domain different from the first domain, the indication (e.g., **810**A, **810**B, **810**C) is a second type of indication different from the first type of indication.

[0293] Providing domain specific indications quickly informs the user (e.g., **820**) whether the digital assistant was able to transmit the instructions for the command (e.g., **824**C) successfully. Associating particular visual, audio, or digital voice indicators with particular domains provide the user quickly with an improved user interface that allows the

user quickly to quickly determine whether more action needs to be taken if the command (e.g., **824**C) was not successfully carried out. For example, when the user quickly says "love that song" a particular tone that is played and associated with the command informs the user (e.g., **820**) quickly that the digital assistant has added the song to a favorites list, without the user quickly having to verify it was added. This reduces the cognitive burden on the user quickly as the association of the success tone with a particular action in a particular domain provides immediate confirmation that the task was carried out. Further, the feedback of the state of the digital assistant allows the user (e.g., **820**) to use the digital assistant more quickly and efficiently.

[0294] In some embodiments, the electronic device (e.g., **800**) is paired (**916**) to the first external device (e.g., **814**, **818**) (e.g., a paired lamp, tv, computer, phone or watch).

[0295] In some embodiments, in accordance with a determination, using the first gaze information, that the set of one or more gaze criteria is not met (e.g., user breaks gaze; user is not looking at an external device for a threshold period of time) for the first external device (e.g., **814**, **818**), the electronic device (e.g., **800**) forgoes transmitting the instruction to transition the first external device (e.g., **818**) from the first state to the second state. In some examples, the digital assistant does not perform the command (e.g., **824**C), when the set of one or more gaze criteria is not met.

[0296] In some embodiments, the set of one or more gaze criteria includes a field of view criterion that is met when a determined offset angle (e.g., **810**) is less than a threshold offset angle (e.g., **804**) (e.g., the maximum acceptable deviation from the user looking directly at the device; a non-zero threshold offset angle; the threshold offset angle is greater than zero degrees), wherein the determined offset angle (e.g., **810**) is an angle formed between: a line-of-sight (e.g., **812**) (e.g., determined from the gaze information) of a gaze of a user, and a calculated line-of-sight (e.g., **808**) from the user (e.g., **820**) to the first external device (e.g., **818**). In some examples, if the first external device (e.g., **818**) is determined to be positioned more than a threshold number of degrees (e.g., outside the field of view (e.g., **806**) of the gaze), the first external device (e.g., **818**) is not in the user's gaze and the instruction is not sent to the first external device (e.g., **818**). In some examples, when the first external device (e.g., **818**) is in the field of view (e.g., **806**) of the gaze by being within the threshold number of degrees, the digital assistant sends the instruction to the first external device (e.g., **818**).

[0297] Using a field of view criterion as a gaze criteria allows the user's gaze to be detected without requiring the user to look directly at the external device (e.g., **818**). Instead, as long as the external device (e.g., **818**) is within a field of view of the user's gaze, the user's gaze satisfies the gaze criteria, allowing the electronic device (e.g., **800**) to determine which external device (e.g., **814**, **818**) is intended. Using the field of view criterion allows the external device (e.g., **818**) to be determined without requiring further interaction with the user, which enhances the operability of the digital assistant and makes the gaze detection by the electronic device (e.g., **800**) more efficient (e.g., the external device (e.g., **818**) is more easily identified since the external device (e.g., **818**) does not have to be in the direct line-of-sight of the user (e.g., **820**)).

[0298] In some embodiments, the line-of-sight (e.g., **812**) of the gaze of the user (e.g., **820**) is based on a head position of the user (e.g., **820**) (e.g., the same user from whom the audio user input request was received). In some examples, the user's head position may be affected by the tilt or rotation (e.g., yaw, pitch, and roll) of the user's head. In some examples, head position tracking is used to determine whether the first external device (e.g., **818**) is in the field of view (e.g., **806**).

[0299] In some embodiments, the set of one or more gaze criteria includes a dwell time criterion that is met for the first external device (e.g., **818**) when a dwell time (e.g., duration) of the gaze determined from the first gaze information is more than a threshold period of time (e.g., a non-zero period of time). In some examples, when the dwell time of the gaze of the user (e.g., **820**) is less than a threshold amount of time.

[0300] In some embodiments, subsequent to the electronic device (e.g., **8000**) transmitting (e.g., to the first external device (e.g., **818**)) the instruction to transition the first external device (e.g., **818**) from the first state to the second state, the electronic device (e.g., **800**) provides an indication (e.g., **810**A, **810**B, **810**C) (e.g., audio or visual) that the first external device (e.g., **818**) is in the second state (e.g., the electronic device (e.g., **800**) provides a visual indication (e.g., **810**A) to indicate that the first external device (e.g., **818**) has turned on). In some examples, the electronic device (e.g., **800**) provides a different indication (e.g., **810**A, **810**B, **810**C) when the command (e.g., **824**C) was not able to be executed (e.g., the first external device (e.g., **818**) is still in the first state).

[0301] In some embodiments, while the first external device (e.g., **818**) is in the first state and a second external device is in a third state and in accordance with a determination (and, optionally, in response to receiving the audio user input request), using the first gaze information, that the set of one or more gaze criteria is met for the second external device (e.g., appliance such as a lamp, tv, or computer): the electronic device (e.g., **800**) transmits (e.g., to the first external device (e.g., **818**)), based on the first command (e.g., **824**C), an instruction to transition the second external device (e.g., **814**) from the third state to a fourth state (e.g., the electronic device (e.g., **800**) performs an action on the second external device (e.g., **814**) by transitioning the second external device from an ON state to an OFF state or from an OFF state to an ON state).

[0302] In accordance with a determination (and, optionally, in response to receiving the audio user input request), using the first gaze information, that the set of one or more gaze criteria is not met for the second external device (e.g., **814**) (e.g., appliance such as a lamp, tv, or computer): the electronic device (e.g., **800**) forgoes transmitting (e.g., to the first external device), based on the first command (e.g., **824**C), the instruction to transition the second external device (e.g., **814**) from the third state to the fourth state (e.g., the electronic device (e.g., **800**) forgoes performing an action on the second external device, such as forgoing transitioning the second external device from an ON state to an OFF state or from an OFF state to an ON state).

[0303] In some embodiments, while the first external device (e.g., **818**) is in the second state (e.g., the first external device **9** e.g., **818**) may also be in the first state) and a second external device (e.g., **814**) is in a third state (**918**): the electronic device (e.g., **800**) receives (**920**) a second user input request (e.g., a user speech input) including a second command (e.g., **826**C).

[0304] The electronic device (e.g., **800**) obtains (**922**), using one or more camera sensors (e.g., **816**A, **816**B), second gaze information (e.g., to transition the second external device (e.g., **814**) to a second state, telling the digital assistant to turn on "that" light).

[0305] In accordance with a determination, using the second gaze information, that the set of one or more gaze criteria is met for the second external device (e.g., **814**) (e.g., the second gaze identifies a second external device (e.g., **814**) by satisfying gaze criteria such as the angle of the gaze being within a threshold number of degrees, gaze is in a particular direction, duration is greater than or equal to a predetermined time period): the electronic device (e.g., **800**) transmits (**924**), based on the second command (e.g., **824**C), an instruction to transition the second external device (e.g., **814**) from the third state to a fourth state (e.g., changing the state of the second external device (e.g. **814**): the floor lamp turned on) while the first external device (e.g., **818**) remains in the second state (e.g., without changing the state of the first external device: the table lamp remains on).

[0306] In some embodiments, while the second external device (e.g., **814**) is in the third state and in accordance with a determination, using the second gaze information, that the set of one or more gaze criteria is not met for the second external device (e.g., **814**) (e.g., second external device is not identified): the electronic device (e.g., **800**) forgoes (**926**) transmitting the instruction to transition the second external device (e.g., **814**) from the third state to the fourth state (e.g., second external device (e.g., **814**) remains in the same state as before).

[0307] In some embodiments, subsequent to transmitting (e.g., to the second external device (e.g., **814**) the instruction to transition the second external device (e.g., **814**) from the third state to the fourth state, the electronic device (e.g., **800**) provides an indication (e.g., **810**A, **810**B, **810**C) (e.g., visual or audio indication) that indicates that the second external device (e.g., **814**) is in the fourth state (e.g., provide a visual (e.g., **810**A) or audio (e.g., **810**B) indication to indicate the second external device (e.g., **814**) has changed state). In some examples, the electronic device (e.g., **800**) provides a different indication (e.g., **810**A, **810**B, **810**C) when the second command (e.g., **824**C) was not able to be executed, (e.g., the second external device (e.g., **813**) is still in the third state).

[0308] In some embodiments, subsequent to transmitting (e.g., to the first external device (e.g., **818**)) the instruction to transition the first external device (e.g., **818**) from the first state to the second state: the electronic device (e.g., **800**) receives from the third external device (e.g. a watch or a phone), an indication of an input received by the third external device (e.g., **830**) (e.g., rotation of a crown (e.g., **832**) on a third external device). The electronic device (e.g., **800**) transmits a second instruction, to the first external device (e.g., **818**) based on a recency of the electronic device (e.g., **800**) having instructed the first external device (e.g. **818**) (e.g., a third instruction to transition the first external device (e.g., **818**) (e.g., dim the brightness of the light of the table lamp after the table lamp has been turned on). In some examples, third external device (e.g., **830**) doesn't know which external device (e.g., **814**, **818**) the third external device (e.g., **840**) is interacting with. The third external device (e.g., **830**) only receives the input and transmits the user input to the electronic device (e.g., **800**), which is the device that figures out which external device (e.g., **814**, **818**) to send the instruction to (e.g., dim the light).

[0309] Using a third external device (e.g., **830**) to further control the most recent external device (e.g., **818**) that the electronic device (e.g., **800**) sent an instruction to, provides the user (e.g., **820**) with additional user interfaces to have more control over the external device (e.g., **818**).

[0310] Providing a user (e.g., **820**) with additional user interfaces and mechanisms for control devices without having to go through the digital assistant provides more efficient user interfaces and reduces the number of interactions needed with the digital assistant to perform a command. Rather than having to interact with the digital assistant to process additional commands, the user (e.g., **820**) may simply use an external device (e.g., **830**) to provide more precise input that is sometimes difficult to explain in words. Receiving additional user input based on input received from a third external device (e.g., **840**), enhances the operability of the digital assistant and makes the user interface of controlling external devices (e.g., **818**) more efficient as additional commands are not needed.

[0311] Note that details of the processes described above with respect to method **900** (e.g., FIGS. **9**A-**9**B are also applicable in an analogous manner to the methods described below/above. For example, method **900** optionally includes one or more of the characteristics of the various methods described above with reference to method **700** and below with reference to method **1100**. For example, method **900** may include one or more of the characteristics of the various methods described above and below with reference to the processes in FIGS. **7**A, **7**B, **11**A, and **11**B. For brevity, these details are not repeated below.

[0312] It should be understood that the particular order in which the operations in the following FIGS. **7**A-**7**B and **11**A-**11**B have been described is exemplary and not intended to indicate that the described order is the only order in which the operations could be performed. One of ordinary skill in the art would recognize various ways to reorder the operations described herein, as well as excluding certain operations. For brevity, these details are not repeated here. Additionally, it should be noted that aspects of the methods and processes described throughout this description may be incorporated with one another.

[0313] FIGS. **10**A-**10**D illustrate exemplary user interfaces for providing different indicators to indicate the digital assistant's recognition of various users (e.g., **1020**, **1030**) speaking commands in a room, once the digital assistant has been activated, in accordance with some embodiments. The digital assistant is activated based on a trigger word, a gaze as discussed above with respect to method **700**, a button press, a wrist raise while the user is wearing a wearable electronic device (e.g., a watch) (e.g., FIGS. **6**E-**6**G), and/or the user pointing an external device (e.g., a phone) at electronic device **1000**. In some examples, the electronic device **1000** activates the digital assistant in response to determining that a user (e.g., **1020**, **1030**) is looking at electronic device **1000** (e.g., in accordance with a determination that a set of one or more activation criteria are met using gaze information, as discussed above) and receives commands once the digital assistant is activated. The techniques illustrated in FIGS. **10**A-**10**D also optionally include aspects of method **900** and the techniques described with respect to FIGS. **8**A-**8**L, which illustrate an electronic device using the user's gaze to determine a particular external

object (e.g., table lamp **1018**, floor lamp) associated with the command. The techniques in these figures are used to illustrate the processes described below, including the processes in FIGS. **11**A-**11**B.

[0314] As illustrated in FIGS. **10**A-**10**D, users Jane **1030** and Jack **1020** are in the same room as electronic device **1000**. In FIG. **10**A, the digital assistant of electronic device is activated, such as using aspects of the techniques described above. Electronic device **1000** receives a spoken command (e.g., **1024**A, **1024**B) from one or more users, (e.g., Jane **1030** and Jack **1020**), to control various external devices, such a floor lamp **1014** or a table lamp **1018** in the room. In response to electronic device **1000** receiving the command (e.g., **1024**A, **1024**B), electronic device **1000** determines a user identity (e.g., **1020**, **1030**) of the particular user who spoke the command (e.g., **1024**A, **1024**B). In some embodiments, an external device (e.g., external cameras **1016**A, a phone, or integrated camera **1016**B) determines the user identity corresponding to the command, and sends the user identity to electronic device **1000**. In some embodiments, electronic device **1000** obtains information about the user speaking using an external or internal device (e.g., external cameras **1016**A, a phone, or integrated camera **1016**B). In some embodiments, the electronic device **1000** uses the obtained information to perform facial recognition, voice recognition, or to calculate a distance to the user speaking to determine a user identity.

[0315] In some embodiments, prior to the electronic device activating the digital assistant, users Jane **1030** and Jack **1020** have optionally registered user profiles that are stored on electronic device **1000**. In some embodiments, the user profiles are stored on a server and are accessible by electronic device **1000**. The user profiles include settings and preferences (e.g., type of accent for the digital assistant voice, the gender for the digital assistant voice, indicator tones, LED light colors) that the electronic device may use to customize indicators **1010**A, **1010**B, **1010**C used by the electronic device to indicate when the digital assistant is activated or when a command **1024**A, **1024**B has been successfully or unsuccessfully performed. For example, the user may customize the digital indicator settings by providing a value for each characteristic of the indicator. In some examples, the characteristics of the indicator include the type (e.g., visual, audio, voice confirmation or a combination of any of the types), a light color for a visual indicator, a tone for an audio indicator, and customizations (e.g., language, type of accent and gender) for the digital voice. The user profile also optionally stores information about the user provided by users Jane **1030** and Jack **1020** during registration with the user's consent, such as the user's name, birthdate, gender, and/or address. The user profile settings and preferences may be updated at any time. The techniques for user registration on electronic device **1000** described here may be used above in combination with methods **700** and **900** with reference to the processes in FIGS. **7**A, **7**B, **9**A, and **9**B. For brevity, these details are not repeated above.

[0316] In FIG. **10**A, electronic device **1000** receives command **1024**A, "turn on the table lamp," from Jane **1030** when the digital assistant is activated. As electronic device **100** is receiving the command **1024**A being spoken by Jane **1030** (or after receiving the command **1024**A spoken by Jane **1030**), electronic device **1000** determines that the speaker is Jane **1030**. For example, device **1000** uses the external cameras **1016**A, integrated camera **1016**B, and/or microphones, determines that the received command **1024**A is coming from the direction in which Jane **1030** is sitting. Accordingly, electronic device **1000** displays a light pattern **1010**D (e.g., three illuminated dots) on display **1002** in a position that corresponds to the direction in which Jane **1030** is sitting. In some embodiments, electronic device **1000** displays an animated light pattern **1010**D on display **1002**, where the light pattern is animated to point to the direction in which Jane **1030** is sitting.

[0317] In FIG. **10**A, when the digital assistant determines that the speaker of command **1024**A is Jane **1030**, electronic device **1000** updates a value of a characteristic (e.g., a color of a light, a particular sound, words, or a specific digital assistant voice) of an indicator (e.g., visual **1010**A, audio **1010**B, or a digital assistant confirmation **1010**C) to a value corresponding to Jane **1030** (e.g., a purple light, a two-note tone, Jane's name, a British accent, female voice). In some embodiments, the value of the characteristic is a value obtained from Jane's user profile. For example, the electronic device sets the color of light pattern **1010**D (e.g., three illuminated dots) on display **1002** to be a color (such as purple) that corresponds to Jane **1030**. As a result, electronic device **1000** indicates to Jane **1030** and Jack **1020** that device **1000** has recognized the speaker as Jane **1030** and that one or more of the preferences of Jane **1030** will be used.

[0318] In contrast, in FIG. **10**C when the digital assistant determines that the speaker of command **1024**B is Jack **1020**, who is a different user than Jane **1030**, electronic device **1000** updates the value of the characteristic (e.g., a color of a light, a sound, or a digital assistant voice) of the indicator (e.g., a visual indicator or an audio indicator) to a value corresponding to Jack **1020** (e.g., a blue light, a one-note tone, Jack's name, an Australian accent, male voice), where the value of the characteristic associated with Jack **1020** is different form the value of the characteristic associated with Jane **1030**. For example, the electronic device sets the color of light pattern **1010**E (e.g., three illuminated dots) on display **1002** to be a color (such as blue) that corresponds to Jack **1020**. As a result, electronic device **1000** indicates to Jane **1030** and Jack **1020** that device **1000** has recognized the speaker as Jack **1020** and that one or more of the preferences of Jack **1020** will be used.

[0319] In some embodiments, Jack **1020** and Jane **1030** may have values of characteristics associated with each person that are the same. For example, Jack **1020** and Jane **1030** may both choose the American, female, digital assistant voice. In some embodiments, the values of the characteristic of the indicators associated with each user is unique. For example Jane **1030** and Jack **1020** may have different light colors (e.g., purple for Jane vs. blue for Jack) and different sounds (e.g., two note sound vs. one note sound) associated with the respective user.

[0320] In FIG. **10**B, once electronic device **1000** has identified that Jane **1030** is speaking the command **1024**A, electronic device **1000** determines whether a set of one or more performance criteria are met, as discussed above with respect to method **900**. The set of one or more performance criteria includes one or more of: a completion criterion, an actionable criterion, and an activation criteria. When the set of one or more performance criteria are met, electronic device **1000** performs the action, such as transmitting the instructions to floor lamp **1014** to be turned on.

[0321] When electronic device **1000** has performed command **1024**A, electronic device **1000** optionally provides an indicator (e.g., indicator **1010**A, **1010**B, **1010**C) using Jane's **1030** preferences to indicate that the command has been successfully performed. The types of indicators (e.g., visual, audio, and digital voice confirmations) and their characteristics are discussed above with respect to method **700** and **900**. In FIG. **10**B, electronic device **1000** has determined that the speaker is Jane **1030** and, based on that determination, provides a confirmation using a female digital assistant voice confirmation with an American accent, "Jane, the floor lamp has been turned on." The name "Jane" is obtained from Jane's **1030** user profile, which Jane **1030** provided during registration. Electronic device **1000** uses a female American accent based on a previous preference stored as part of Jane's **1030** user profile. A purple light pattern may be displayed as the visual indicator on electronic device **1000** to indicate that the digital assistant has processed Jane's command **1024**A. A tone associated with Jane **1030** may also be played as the audio indicator on electronic device **1000** to indicate that the digital assistant has processed Jane's command **1024**A. As a result, both Jane **1030** and Jack **1020** recognize that Jane was identified as the speaker.

[0322] FIG. **10**C similarly illustrates Jack **1020** speaking command **1024**B "turn on the floor lamp." Electronic device **1000** determines that the speaker is Jack **1020** using, for example, external cameras **1016**A, integrated camera **1016**B, and/or microphones to determine that the command **1024**B is coming from the direction in which Jack **1030** is sitting. In response, electronic device **1000** displays a light pattern **1010**E (e.g., three illuminated dots) on display **1002** in a position that corresponds to the direction in which Jack **1020** is standing. In some embodiments, electronic device **1000** displays an animated light pattern **1010**E on display **1002**, where the light pattern is animated to point to the direction in which Jack **1020** is standing. For example, the electronic device sets the color of light pattern **1010**E (e.g., three illuminated dots) on display **1002** to be a color (such as blue) that corresponds to Jack **1020**. As a result, electronic device **1000** indicates to Jane **1030** and Jack **1020** that device **1000** has recognized the speaker as Jack **1020** and that one or more of the preferences of Jack **1020** will be used.

[0323] In FIG. **10**D, once electronic device **1000** has identified that Jack **1020** is speaking the command **1024**B, electronic device **1000** determines whether a set of one or more performance criteria are met, as discussed above with respect to method **900**. The set of one or more performance criteria includes one or more of: a completion criterion, an actionable criterion, and an activation criteria. When the set of one or more performance criteria are met, electronic device **1000** performs the action, such as transmitting the instructions to floor lamp **1014** to be turned on.

[0324] When electronic device **1000** has performed command **1024**B, electronic device **1000** optionally provides an indicator (e.g., indicator **1010**A, **1010**B, **1010**C) using Jacks' **1020** preferences to indicate that the command has been successfully performed. The types of indicators (e.g., visual, audio, and digital voice confirmations) and their characteristics are discussed above with respect to method **700** and **900**. In FIG. **10**D, electronic device **1000** has determined that the speaker is Jack **1020** and, based on that determination, provides a confirmation using a male digital assistant voice with an Australian accent, "Jack, the table lamp has been turned on." The name "Jack" is obtained from Jack's **1030** user profile, which Jack **1020** provided during registration. Electronic device **1000** uses a male Australian accent based on a previous preference stored as part of Jack's **1020** user profile. A blue light pattern may be displayed as the visual indicator on electronic device **1000** to indicate that the digital assistant has processed Jack's command **1024**B. The electronic device optionally plays a custom tone associated with Jack **1020** as the audio indicator to indicate that the digital assistant has processed Jack's command **1024**B.

[0325] FIGS. **11**A-**11**B are a flow diagrams illustrating a method **1100** using an electronic device in accordance with some embodiments. Method **1100** is performed at an electronic device (e.g., **100**, **300**, **500**, **600**, **640**, **800**, **1000**). Some operations in method **1100** are, optionally, combined, the orders of some operations are, optionally, changed, and some operations are, optionally, omitted.

[0326] As described below, method **1100** provides an intuitive way for the digital assistant to show recognition of different users (e.g., **1020**, **1030**) by providing unique indicator for each person. The method reduces the cognitive burden on a user (e.g., **1020**, **1030**) because the user is able to confirm that the digital assistant processed the command of a particular user when multiple users are speaking commands in the room. Therefore, the digital assistant is able to provide immediate feedback to indicate which user's commands were processed.

[0327] After the digital assistant on the electronic device (e.g., **1000**) has been activated based on a trigger word, gaze, and/or button press, electronic device **1000** receives (**1106**) an audio user input request (e.g., **1024**A, **1024**B). In some examples, in response to receiving the audio user input request (e.g., **1024**A, **1024**B), the electronic device (e.g., **1000**) determines a user identity (e.g., a particular user (e.g., **1020**, **1030**)) corresponding to the audio user input request (e.g., **1024**A, **1024**B). In some examples, a remote device (e.g., cameras **1016**) determines the user identity corresponding to the audio user input request (e.g., **1024**A, **1024**B) and the electronic device (e.g., **1000**) receives results of the determination.

[0328] In some embodiments, prior to the electronic device (e.g., **1000**) receiving the audio user input request (e.g., **1024**A, **1024**B), the electronic device (e.g., **1000**) receives (**1102**) registration information (e.g., user profile information). In some embodiments, the electronic device (e.g., **1000**) associates (**1104**), using the registration information, a first value of a characteristic (e.g., a light, a sound, a digital assistant voice) of an indicator (e.g., **1010**A, **1010**B, **1010**C) of electronic device (e.g., **1000**) with a first user (e.g., **1020**) (and not the second user) and a second value of the characteristic of an indicator of electronic device (e.g., **1000**) with a second user (e.g., **1030**) (and not the first user). In some examples, during a registration process, the electronic device (e.g., **1000**) receives a name of a user and later associates the name with the voice of the user (e.g., **1020**, **1030**). In some examples, during a registration process, the electronic device (e.g., **1000**) associates a color of a visual indicator (e.g., LED color) to the user (e.g., **1020**, **1030**).

[0329] Associating values of a characteristic of the indicator (e.g., **1010**A, **1010**B, **1010**C) with preferences from the user's user profile, enables the digital assistant to provide feedback to confirm the digital assistant has identified the

specific user (e.g., **1020**, **1030**) speaking the command. Providing customized feedback to confirm the user (e.g., **1020**, **1030**) enhances the operability of the digital assistant and makes the user-device interface more efficient (e.g., the user (e.g., **1020**, **1030**) knows the digital assistant has processed the user's command (e.g., **1024**A, **1024**B)) by providing confirmation to the user (e.g., **1020**, **1030**) using the user's preferences in the indicator (e.g., **1010**A, **1010**B, **1010**C). This reduces the cognitive burden on the user (e.g., **1020**, **1030**) as the association of the success indicator with a particular user's settings provides immediate confirmation that the task was carried out and reduces the confusion when there are multiple users speaking simultaneously or other speakers in the background. Further, the feedback of the state of the digital assistant allows the user (e.g., **1020**, **1030**) to use the digital assistant more quickly and efficiently.

[0330] The electronic device receives (**1106**) an audio user input request.

[0331] In some embodiments, the audio user input request (e.g., **1024**A, **1024**B) is a user speech input (e.g., utterance, a command) and the updating of (**1108**) the value of the characteristic (e.g., a light, a sound, a digital assistant voice) is not based on transcribed contents of the audio user input request (e.g., **1024**A, **1024**B) (e.g., the value of the characteristic is not based on the words said by the user, but instead is based on the identity of the user).

[0332] In accordance with a determination that the audio user input request corresponds to a first user (e.g., **1020**) (and, optionally, in response to receiving the audio user input request), the electronic device (e.g., **1000**) updates (**1110**) a value of a characteristic (e.g., a color of a light, a sound, or a digital assistant voice) of an indicator (e.g., **1010**A, **1010**B, **1010**C) (e.g., a visual indicator, an audio indicator, or a digital assistant voice confirmation) to a first value corresponding to the first user (e.g., **1020**) (e.g., the value of the characteristic is associated with the identity of the first user).

[0333] In accordance with a determination that the audio user input request corresponds to a second user (e.g., **1030**) (and, optionally, in response to receiving the audio user input request), different from the first user, electronic device (**1000**) updates (**1112**) the value of the characteristic (e.g., a color of a light, a sound, or a digital assistant voice) of the indicator (e.g., **1010**A, **1010**B, **1010**C) (e.g., a visual indicator, an audio indicator, or a digital assistant voice confirmation) to a second value corresponding to the second user (e.g., **1030**) (e.g., the value of the characteristic is associated with the identity of the second user), the second value being different from the first value.

[0334] The electronic device (e.g., **1000**) responds (**1114**) to the audio user input request (e.g., **1024**A, **1024**B) using the indicator (e.g., **1010**A, **1010**B, **1010**C) (e.g., the electronic device (e.g., **1000**) displays a light, plays a sound, using a particular digital assistant voice, or using the user's name in the response), wherein the indicator (e.g., **1010**A, **1010**B, **1010**C) includes the updated value of the characteristic (e.g., a color of a light, a sound, or a digital assistant voice).

[0335] In some embodiments, the electronic device (e.g., **1000**) responding to the audio user input request (e.g., **1024**A, **1024**B) using the indicator (e.g., **1010**A, **1010**B, **1010**C) includes the electronic device (e.g., **1000**) displaying (**1116**) a visual indicator (e.g., **1010**A) (e.g., light) using the value of the characteristic (e.g., light color, position of displayed light) corresponding to the user associated with the audio user input request (e.g., **1024**A, **1024**B).

[0336] In some embodiments, the electronic device (e.g., **1000**) responding to the audio user input request (e.g., **1024**A, **1024**B) using the indicator includes the electronic device (e.g., **1000**) providing (**1118**) an audio indicator (e.g., **1010**B) (e.g., a voice, a sound) using the value of the characteristic (the value of the characteristic could be an accent, gender, pitch, voice, sound, or a particular digital assistant voice, such as American, Female digital assistant voice) corresponding to the user (e.g., **1020**, **1030**) associated with the audio user input request (e.g., **1024**A, **1024**B). In some examples, the user selected the American, Female digital assistant voice in the digital assistant voice settings as indicated in the user profile of the user (e.g., **1020**, **1030**).

[0337] In some embodiments, the indicator is a visual indicator (e.g., **1010**A) and in response to the electronic device (e.g., **1000**) receiving the audio user input request (e.g., **1024**A, **1024**B) and prior to the determination that the audio user input request (e.g., **1024**A, **1024**B) corresponds to a particular user (e.g., **1020**, **1030**), the electronic device (e.g., **1000**) displays the visual indicator (e.g., **1010**A) with a default value (e.g., white light) of the characteristic (e.g., displaying a first light color when the digital assistant is activated). In some embodiments, the default value of the characteristic is different from the first value and the second value.

[0338] In some embodiments, the characteristic of the audio indicator (e.g., **1010**B) (e.g., the digital assistant providing confirmation using words) is a term (e.g., a word, a name of) corresponding to a user associated with (e.g., that provides) the audio user input request (e.g., **1024**A, **1024**B).

[0339] Using the name of the user (e.g., **1020**, **1030**) in the digital assistant voice confirmation indicator (e.g., **1010**C) enables the digital assistant to provide feedback to confirm the digital assistant has identified the specific user (e.g., **1020**, **1030**) speaking the command. Providing customized feedback to confirm the user (e.g., **1020**, **1030**) enhances the operability of the digital assistant and makes the user-device interface more efficient (e.g., the specific user (e.g., **1020**, **1030**) knows the digital assistant has processed the user's command (e.g., **1024**A, **1024**B). This reduces the cognitive burden on the user (e.g., **1020**, **1030**) as the use of the user's name provides immediate confirmation that the task was carried out and reduces the confusion when there are multiple users (e.g., **1020**, **1030**) speaking simultaneously or other speakers in the background. Further, the feedback of the state of the digital assistant allows the user (e.g., **1020**, **1030**) to use the digital assistant more quickly and efficiently.

[0340] In some embodiments, the electronic device (e.g., **1000**) determines a user associated with the audio user input request based on one or more of voice recognition, facial recognition, and a direction of the audio user input request.

[0341] The foregoing description, for purpose of explanation, has been described with reference to specific embodiments. However, the illustrative discussions above are not intended to be exhaustive or to limit the invention to the precise forms disclosed. Many modifications and variations are possible in view of the above teachings. The embodiments were chosen and described in order to best explain the principles of the techniques and their practical applications. Others skilled in the art are thereby enabled to best utilize

the techniques and various embodiments with various modifications as are suited to the particular use contemplated.

[0342] Although the disclosure and examples have been fully described with reference to the accompanying drawings, it is to be noted that various changes and modifications will become apparent to those skilled in the art. Such changes and modifications are to be understood as being included within the scope of the disclosure and examples as defined by the claims.

[0343] As described above, one aspect of the present technology is the gathering and use of data available from various sources to better enable users to control electronic devices. The present disclosure contemplates that in some instances, this gathered data may include personal information data that uniquely identifies or can be used to contact or locate a specific person. Such personal information data can include demographic data, location-based data, telephone numbers, email addresses, twitter IDs, home addresses, data or records relating to a user's health or level of fitness (e.g., vital signs measurements, medication information, exercise information), date of birth, or any other identifying or personal information.

[0344] The present disclosure recognizes that the use of such personal information data, in the present technology, can be used to the benefit of users. For example, the personal information data can be used to personalize device interactions. Accordingly, use of such personal information data enables users to more easily control electronic devices. Further, other uses for personal information data that benefit the user are also contemplated by the present disclosure. For instance, health and fitness data may be used to provide insights into a user's general wellness, or may be used as positive feedback to individuals using technology to pursue wellness goals.

[0345] The present disclosure contemplates that the entities responsible for the collection, analysis, disclosure, transfer, storage, or other use of such personal information data will comply with well-established privacy policies and/or privacy practices. In particular, such entities should implement and consistently use privacy policies and practices that are generally recognized as meeting or exceeding industry or governmental requirements for maintaining personal information data private and secure. Such policies should be easily accessible by users, and should be updated as the collection and/or use of data changes. Personal information from users should be collected for legitimate and reasonable uses of the entity and not shared or sold outside of those legitimate uses. Further, such collection/sharing should occur after receiving the informed consent of the users. Additionally, such entities should consider taking any needed steps for safeguarding and securing access to such personal information data and ensuring that others with access to the personal information data adhere to their privacy policies and procedures. Further, such entities can subject themselves to evaluation by third parties to certify their adherence to widely accepted privacy policies and practices. In addition, policies and practices should be adapted for the particular types of personal information data being collected and/or accessed and adapted to applicable laws and standards, including jurisdiction-specific considerations. For instance, in the US, collection of or access to certain health data may be governed by federal and/or state laws, such as the Health Insurance Portability and Accountability Act (HIPAA); whereas health data in other countries may be subject to other regulations and policies and should be handled accordingly. Hence different privacy practices should be maintained for different personal data types in each country.

[0346] Despite the foregoing, the present disclosure also contemplates embodiments in which users selectively block the use of, or access to, personal information data. That is, the present disclosure contemplates that hardware and/or software elements can be provided to prevent or block access to such personal information data. For example, in the case of digital assistants, user gaze, and user profiles, the present technology can be configured to allow users to select to "opt in" or "opt out" of participation in the collection of personal information data during registration for services or anytime thereafter. In another example, users can select not to provide personal information for user profiles. In yet another example, users can select to limit the length of time personal information is maintained or entirely prohibit the development of a baseline profile. In addition to providing "opt in" and "opt out" options, the present disclosure contemplates providing notifications relating to the access or use of personal information. For instance, a user may be notified upon downloading an app (e.g., a digital assistant app) that their personal information data will be accessed and then reminded again just before personal information data is accessed by the app.

[0347] Moreover, it is the intent of the present disclosure that personal information data should be managed and handled in a way to minimize risks of unintentional or unauthorized access or use. Risk can be minimized by limiting the collection of data and deleting data once it is no longer needed. In addition, and when applicable, including in certain health related applications, data de-identification can be used to protect a user's privacy. De-identification may be facilitated, when appropriate, by removing specific identifiers (e.g., date of birth, etc.), controlling the amount or specificity of data stored (e.g., collecting location data a city level rather than at an address level), controlling how data is stored (e.g., aggregating data across users), and/or other methods.

[0348] Therefore, although the present disclosure broadly covers use of personal information data to implement one or more various disclosed embodiments, the present disclosure also contemplates that the various embodiments can also be implemented without the need for accessing such personal information data. That is, the various embodiments of the present technology are not rendered inoperable due to the lack of all or a portion of such personal information data. For example, users can interact with electronic devices based on non-personal information data or a bare minimum amount of personal information.

What is claimed is:

1. An electronic device, comprising:

one or more processors; and

memory storing one or more programs configured to be executed by the one or more processors, the one or more programs including instructions for:

while a digital assistant of the electronic device is not activated:

obtaining, using one or more camera sensors, first gaze information; and

in accordance with a determination that the first gaze information satisfies a set of one or more activation criteria:

activating the digital assistant of the electronic device; and

providing an indication that the set of one or more activation criteria has been satisfied.

2. The electronic device of claim **1**, the one or more programs further including instructions for:

in accordance with a determination that the first gaze information does not satisfy the set of one or more activation criteria:

forgo activating the digital assistant.

3. The electronic device of claim **1**, the one or more programs further including instructions for:

receiving an audio user input request to perform a command; and

in response to receiving the audio user input request to perform the command:

in accordance with a determination that a set of one or more performance criteria is met, performing the command based on a content of the audio user input request.

4. The electronic device of claim **3**, wherein the electronic device starts receiving the audio user input request while the digital assistant is not activated.

5. The electronic device of claim **3**, the one or more programs further including instructions for:

producing an audio output, wherein the audio output is being produced at a first volume prior to receiving the audio user input request; and

in response to beginning to receive the audio user input request, reducing a volume of the audio output to a second volume that is lower than the first volume, wherein the second volume is based on a distance between a user and the electronic device.

6. The electronic device of claim **3**, the one or more programs further including instructions for:

in response to receiving the audio user input request to perform the command:

in accordance with a determination that the set of one or more performance criteria is not met, forgo performing the command based on the content of the audio user input request, wherein the set of one or more performance criteria includes a continuous gaze criterion that is met when the electronic device determines that second gaze information indicates that the user does not break a gaze directed at the electronic device for more than a threshold duration between activation of the digital assistant and termination of the audio user input request.

7. The electronic device of claim **3**, the one or more programs further including instructions for:

in response to receiving the audio user input request to perform the command:

in accordance with a determination that the set of one or more performance criteria is not met, providing a headshake indication, wherein the headshake indication includes:

displaying a first pattern of lights on a first side of a display of the electronic device,

subsequent to displaying the first pattern of lights, displaying a second pattern of lights on a second side of the display of the electronic device, and

subsequent to displaying the second pattern of lights, displaying the first pattern of lights on the first side of the display of the electronic device.

8. The electronic device of claim **1**, wherein the indication is a visual indication that indicates that the digital assistant is activated.

9. The electronic device of claim **1**, wherein the indication is an audio indication that indicates that the digital assistant is activated.

10. The electronic device of claim **1**, wherein the set of one or more activation criteria includes a field of view criterion that is met when a determined offset angle is less than a threshold offset angle, wherein the determined offset angle is an angle formed between:

a line-of-sight of a gaze of a user, and

a calculated line-of-sight from the user to the electronic device.

11. The electronic device of claim **10**, wherein the line-of-sight of the gaze of the user is based on a head position of the user.

12. The electronic device of claim **1**, wherein the set of one or more activation criteria includes a dwell time criterion that is satisfied when a dwell time of the gaze determined from the first gaze information is more than a threshold period of time.

13. A non-transitory computer-readable storage medium storing one or more programs configured to be executed by one or more processors of an electronic device, the one or more programs including instructions for:

while a digital assistant of the electronic device is not activated:

obtaining, using one or more camera sensors, first gaze information; and

in accordance with a determination that the first gaze information satisfies a set of one or more activation criteria:

activating the digital assistant of the electronic device; and

providing an indication that the set of one or more activation criteria has been satisfied.

14. A method, comprising:

at an electronic device:

while a digital assistant of the electronic device is not activated:

obtaining, using one or more camera sensors, first gaze information; and

in accordance with a determination that the first gaze information satisfies a set of one or more activation criteria:

activating the digital assistant of the electronic device; and

providing an indication that the set of one or more activation criteria has been satisfied.

* * * * *