

(51) International Patent Classification:
H04B 1/40 (2006.01) **G06F 3/048** (2006.01)(21) International Application Number:
PCT/US2010/050270(22) International Filing Date:
24 September 2010 (24.09.2010)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
12/570,945 30 September 2009 (30.09.2009) US(71) Applicant (for all designated States except US): **T-MOBILE USA, INC.** [US/US]; 12920 SE 38th Street, Bellevue, WA 98006-1350 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **VANCE, Michael, Steffen** [US/US]; c/o T-mobile USA, Inc., 12920 SE 38th Street, Bellevue, WA 98006-1350 (US). **GOULART, Valerie, Lynn** [US/US]; c/o T-mobile USA, Inc., 12920 SE 38th Street, Bellevue, WA 98006-1350 (US). **MANN, Jonathan, L.** [US/US]; c/o T-mobile USA, Inc., 12920 SE 38th Street, Bellevue, WA 98006-1350 (US). **PANCHAL, Prarthana, H.** [US/US]; c/o T-mobile USA, Inc., 12920 SE 38th Street, Bellevue, WA 98006-1350 (US). **WOOD, Paula, H.** [US/US]; c/o T-mobile USA, Inc., 12920 SE 38th Street, Bellevue, WA 98006-1350 (US). **TEMPLE, Sinclair, M.** [US/US]; c/o T-mobile USA, Inc., 12920 SE 38th Street, Bellevue, WA 98006-1350 (US).(US). **MORRISON, Andrew, R.** [US/US]; c/o T-mobile USA, Inc., 12920 SE 38th Street, Bellevue, WA 98006-1350 (US). **BIRAND, Gulay, A.** [US/US]; c/o T-mobile USA, Inc., 12920 SE 38th Street, Bellevue, WA 98006-1350 (US).(74) Agents: **PECK, Robert, C.** et al.; Lee & Hayes, PLLC, 601 W, Riverside Ave. Suite 1400, Spokane, WA 99201 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: UNIFIED INTERFACE AND ROUTING MODULE FOR HANDLING AUDIO INPUT

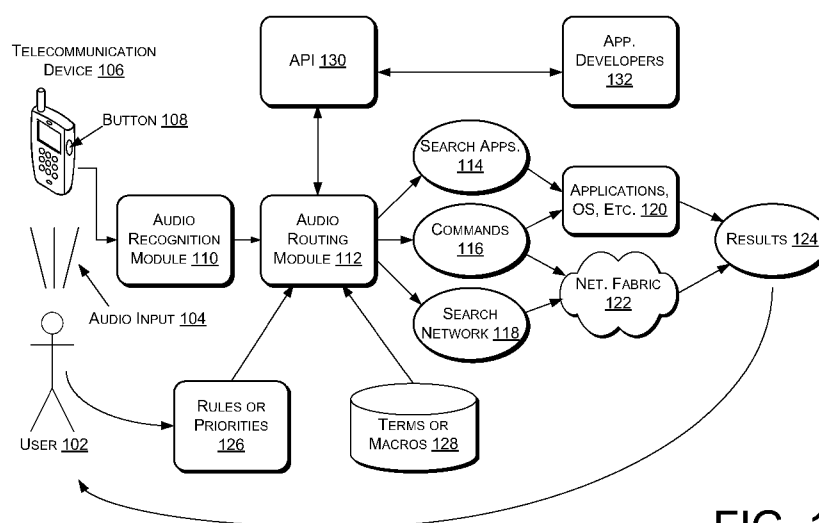


FIG. 1

(57) Abstract: Methods, systems and articles for receiving, by a telecommunication device, audio input through a unified audio interface are disclosed herein. The telecommunication device is further configured to determine whether to handle the audio input as a command, as a request of a telecommunication device application, or as a network search request, the determining being based on one or more user-configurable rules or priorities.



Published:

- *without international search report and to be republished
upon receipt of that report (Rule 48.2(g))*

UNIFIED INTERFACE AND ROUTING MODULE FOR HANDLING AUDIO INPUT**CROSS REFERENCE TO RELATED PATENT APPLICATIONS**

This application claims priority from US Patent Application No. 12/570,945 filed on 30 September 2009, entitled “UNIFIED INTERFACE AND ROUTING MODULE FOR HANDLING AUDIO INPUT,” which is hereby incorporated in its entirety by reference.

BACKGROUND

[0001] Over the past thirty years, telephony has significantly evolved from the once-ubiquitous public switched telephone network service. Advances in technology have added an ever increasing array of features and capabilities to telecommunication devices, such as touch screens, video and still cameras, web browsing capabilities, email sending and receiving capabilities, music and video download, storing and playback capabilities, calendar and contact managing capabilities, global positioning system (GPS) location and navigation capabilities, game playing capabilities, and television capabilities, to name a few. Many of these features and capabilities are provided through specialized applications resident on the telecommunication devices. For example, many telecommunication devices allow the user to further customize the telecommunication device through custom configuration options or by adding third-party software. Such software is available for many different functions, not limited to communications.

[0002] In addition, traditional components have been utilized to provide new capabilities. The microphones of telecommunication devices, for example, are now used for capturing a user's audio input outside the context of a phone call between the

user and another person or system. Applications, such as those mentioned above, accept and process audio input provided through a telecommunication device microphone. Such applications may utilize audio input to create recorded audio notes, voice annotations of a picture, or search queries to be executed on a remote search engine. While these numerous applications all utilize audio input, each provides a different interface and requires different user inputs to initiate the capture of audio input. The captured audio input is then processed by the application triggering the audio capture and is utilized only for the purposes of that application

BRIEF DESCRIPTION OF THE DRAWINGS

[0003] The detailed description is set forth with reference to the accompanying figures, in which the left-most digit of a reference number identifies the figure in which the reference number first appears. The use of the same reference numbers in different figures indicates similar or identical items or features.

[0004] FIG. 1 illustrates an overview of a telecommunication device configured with a unified audio interface and audio routing module, in accordance with various embodiments.

[0005] FIG. 2 illustrates a network view showing a telecommunication device in communication with a remote server having audio recognition capabilities, in accordance with various embodiments.

[0006] FIG. 3 illustrates a component level view of a telecommunication device, in accordance with various embodiments.

[0007] FIG. 4 illustrates a flowchart showing a method of receiving and handling audio input using a telecommunication device, in accordance with various embodiments.

[0008] FIG. 5 illustrates a flowchart showing a method of receiving and handling audio input using a telecommunication device while the telecommunication device is engaged in communication with another device, in accordance with various embodiments.

[0009] FIG. 6 illustrates example user interfaces of a telecommunication device, each rendering a dedicated key, in accordance with various embodiments.

DETAILED DESCRIPTION

[0010] Described herein are telecommunication devices configured with unified audio interfaces and audio routing modules for capturing and handling audio input provided to the telecommunication devices. The unified audio interface includes a dedicated key or button, such as a physical or graphical key, which can be pressed, pressed and held, or otherwise actuated regardless of the execution context of the telecommunication device having the button/key. For example, the button/key may be pressed during an audio communication (e.g., telephone call) or while interacting with an application, such as an email application, a text messaging application, instant messaging application, or any other sort of application or process of the telecommunication device. Upon detecting a press or actuation of the button/key, the unified audio interface captures audio input from a user of the telecommunication device or ambient audio in the area of the telecommunication device and provides the audio input to one or more audio recognition modules, which includes at least a speech recognition component. After audio recognition is performed, the resulting recognized audio input, such as a text translation of speech received from the user of the telecommunication device, is received by an audio routing module of the telecommunication device. The audio routing module evaluates the audio input to

determine whether the audio input should be handled as an audio command, as a request of telecommunication device application(s), or as a network search request.

[0011] In various implementations, the audio routing module utilizes configurable rules or priorities, such as user-configurable or carrier-configurable rules or priorities, in determining how the audio input should be handled. The audio routing module may also or instead rely on a collection of recognized commands which, when encountered within the audio input, indicate that the audio input is to be handled as a command. The audio routing module may also be associated with an application programming interface (API) which enables developers of applications for the telecommunication device to register with the telecommunication device and to provide commands or macros which the audio routing module may utilize in determining how to handle the audio input.

[0012] In some implementations, the unified audio interface detects the press or actuation of the button/key while the telecommunication device is engaged in communication with another device. Upon detecting the press or actuation, the telecommunication device may capture, process, and evaluate the audio input. In evaluating the audio input, the audio routing module of the telecommunication device may perform the determining of how to handle the audio input based at least in part on a context or data associated with the communication with the other device.

[0013] FIG. 1 illustrates an overview of a telecommunication device configured with a unified audio interface and audio routing module. As shown, a telecommunication device 106 receives audio input 104 from a user of the telecommunication device 106. Capture of the audio input 104 may be triggered by the user 102 pressing or actuating a button or key 108 of the telecommunication device 106. Once the audio input 104 is captured, an audio recognition module 110, which includes at least a speech recognition component, may recognize the audio input 104,

performing, for example, audio to text translation. The recognized audio input 104 may then be evaluated by an audio routing module 112, the audio routing module 112 determining whether the recognized audio input 104 should be handled as a request 114 of an application or process of the telecommunication device 106, as a command 116, or as a network search request 118. Based on the evaluating, the audio routing module 112 may route the recognized audio input 104 to a telecommunication device application, OS, process, etc. 120 or to location available via a networking fabric 122 for execution or processing of the recognized audio input 104. Results 124 of the executing or processing may then be provided to the user 102 via the telecommunication device 106.

[0014] In various implementations, the audio routing module 112 performs the determining based on configurable rules or priorities which may be added or modified by the user 102 via a rules or priorities process 126 or by a carrier. Also or instead, the determining may be performed based on commands or macros stored in a term or macro data store 128. In some implementations, commands or macros are registered with the audio routing module 112 via an API 130. The API 130 may be utilized by developers 132 of applications or processes for the telecommunication device 106 to register with or interact with the audio routing module 112.

[0015] As illustrated, the telecommunication device 106 that captures the audio input 104 may be any sort of computing device known in the art. For example, the telecommunication device 106 may be a telecommunication device such as a mobile phone or a PDA, a landline phone, an Internet phones, or may be a media player, a personal computer (PC), a laptop, a netbook, a reading device or any sort of mobile or stationary computing device. As mentioned above, the telecommunication device 106 may have a button or key 108. The button or key 108 may be disposed on a location on a shell or casing of the telecommunication device 106, on a headset associated with

the telecommunication device 106, or rendered on a display of the telecommunication device 106, as well as applications, OS, etc. 120. Also, the telecommunication device 106 may be configured with or store any one or more of the audio recognition module 110, the audio routing module 112, the rules or priorities process 126, the term or macro data store 128, and/or the API 130. An example telecommunication device 106 is illustrated in greater detail by Figure 3 and is described further below.

[0016] In some implementations, the telecommunication device 106 or an associated headset may include a microphone or some other means of capturing audio input 104. The audio input 104 may be any sounds captured by the microphone/capturing means either while the button/key 108 is pressed/actuated and held, during a pre-determined time period following the press/actuation, or following a first press/actuation until a second press/actuation is received. In one implementation, the telecommunication device 106 is configured with a signature/identifier of the user's audio that can be used to filter audio input 104 prior to processing the audio input 104. Such a signature/identifier may have been previously captured using the telecommunication device 106 or another device and may comprise a pre-determined string of words that the user 102 is asked to read aloud. In another implementation, rather than requiring a signature/identifier unique to the user 102, the telecommunication device 106 may be configured with a signature/identifier for a language that can be used to filter the audio input 104 to ensure that the audio input 104 comprises spoken words belonging to a given language. In other implementations, the audio input 104 may be received as an audio file from another computing device, the other computing device acting as a proxy for the user 102.

[0017] In various implementations, the button/key 108 may be a physical control disposed on a casing of the telecommunication device 106. Such a control may be disposed on a front, side, top, bottom, or back of telecommunication device 106 and

may signal that it has been pressed or actuated through traces or by contact with a capacitive surface. In one implementation, the button/key 108 may be one of a plurality of keys of a keyboard. The button/key 108 may include a physical control that is pressed towards the telecommunication devices or a switch that is slid or toggled relative to the plane of the telecommunication device 106 on which the button/key 108 is disposed. In one implementation, the button/key 108 is disposed on a removable, attachable shell which mates with the cover/casing of the telecommunication device 106. In a further implementation, the button/key 108 may be a dedicated button/key associated with initiating an audio capture process. Such a dedicated button/key may not be re-programmable or re-assignable by the user 102. In another implementation, the button/key 108 is disposed on a headset associated with the telecommunication device 106.

[0018] In other implementations, as illustrated further in FIG. 6, the button/key 108 is a graphic user interface (GUI) key which is rendered to the display of the telecommunication device 106, the display being a touch sensitive display which comprises a capacitive surface or a set of light sensors. The GUI key 108 may be rendered on each and every user interface (UI) displayed by the telecommunication device 106, including UIs such as menus, application interfaces, and other interfaces of the telecommunication device 106. In one embodiment, the GUI key 108 is located in the same place relative to the bounds of the telecommunication device display in each displayed UI. Thus, by making the GUI key 108 available on every display, the GUI key 108 is made available regardless of the execution context of the telecommunication device 106.

[0019] The button/key 108 and the associated processes for capturing audio input 108 may comprise a unified audio interface for the capture of any commands, requests, queries, or other interactions that a user 102 may have to be executed or

processed on the telecommunication device 106. Either while the button/key 108 is pressed/actuated and held, between two presses/actuators of the button/key 108, or for a pre-determined time period thereafter, the unified audio interface captures audio input 104. As mentioned above, it may even filter the audio input 104. Once the audio input is captured and stored in temporary storage (such as a buffer) or in permanent storage, the unified audio interface may provide the audio input 104 to one or more audio recognition modules 110 which includes at least a speech recognition module. In another implementation, the unified audio interface may provide the audio input 104 to an audio recognition module 110 in real time, as the audio input 104 is received.

[0020] In various implementations, the audio recognition module 110 may comprise any process or processes capable of receiving audio input 104 in the form of an audio file or buffer and producing a text translation of the audio input 104, the text translation comprising “recognized” audio input 104. The audio recognition module 110 may be a process of the telecommunication device 106, of a remote server (as shown in FIG. 2), of a telecommunication service provider, or of some other remote service provider. In one implementation, the audio recognition module 110 may comprise processes implemented on multiple devices. As mentioned above, in one implementation, the audio recognition module 110 may receive the audio input 104 in real time, as it is provided.

[0021] The audio recognition module 110 may utilize any technique or techniques known in the art to recognize the received audio input 104. In some implementations, the audio recognition module 110 may utilize learning models that make progressively better translations of the audio input 104 based on results of previous translations (e.g., such as whether the user is required to choose between multiple options to disambiguate the audio input 104). The audio recognition module 110 may start with a learning model associated with a language or associated with a

specific group sharing a language and some common trait (e.g., age). The learning model may also be selected and pre-loaded by a carrier, for example, based on a billing zip code of the user 102. In one implementation, the audio recognition module 110 may then refine that learning model to take into account idiosyncratic patterns (i.e., voice characteristics) in the user's speech. The refined learning model can comprise a speaker profile and be associated with a user login or identifier, with a telecommunication services account, or with the telecommunication device 106. If associated with a user login, identifier, or account, then the same refined learning model may be accessed and utilized by multiple telecommunication devices.

[0022] In some implementations, the audio recognition module 110 may facilitate training of the learning model based on inputs of user 102. Either in response to the user 102 selecting a training mode or automatically at pre-determined training opportunities, the audio recognition module 110 may allow the user to speak a number of different phrases and, if necessary, select between disambiguating options. Based on the training, the audio recognition module 110 may refine and update the learning model.

[0023] In further implementations, rather than utilizing a learning model customized for the individual user, the audio recognition module 110 can utilize a group learning model that is refined and updated based on results of translations performed for users of the group.

[0024] In some implementations, the audio recognition module 110 or another process of the telecommunication device 106 may provide potential meanings of the audio input 104 to the user 102 to resolve the meaning of the audio input 104 when the audio input 104 is ambiguous. For example, if the user 102 were to say something that is recognized as "brable", the user could either have said (or be trying to say) "stable" or "table." The audio recognition module 110 or process could provide both of these

and perhaps other suggestions to the user 102 for the user 102 to select from. Such suggestions and selections could be made and received either through a GUI or through further audio output/input. In one implementation, the audio recognition module 110 might recognize a term, but not recognize whether the user 102 intends the term as a command for an application or as a search request. If the user 102 were to say “wallpaper”, the user 102 could mean “change the wallpaper” or “show me wallpapers”, the first being a command and the second being a search request that could be fulfilled by an application or search service. Again, the audio recognition module 110 or process could provide both of these and perhaps other suggestions to the user 102 for the user 102 to select from.

[0025] In further implementations, audio recognition module 110 may perform the disambiguation of the audio input 104 with reference to application or context information, such as contact lists, browsing history, a call log, etc. For example, if user 102 speaks the name “Thorbjorn”, the audio recognition module 110 may have trouble understanding the input (i.e., which “Thorbjorn” is being referred to). To disambiguate the input, the audio recognition module 110 may examine the contact list and note that the name Thorbjorn Janssen is included in the list. The audio recognition module 110 may then either assume that the contact is who/what the user 102 was referring to, or may suggest the contact as one of a plurality of disambiguation suggestions for the user 102 to select from. If multiple Thorbjorns are found in the contact lists, the audio recognition module 110 may examine the call log and guess whichever Thorbjorn has been called the most as the Thorbjorn being referred to.

[0026] In some implementations, as mentioned above, the audio recognition module 110 may receive the audio input 104 in real time, as the user 102 is speaking it. To speed along disambiguation, the audio recognition module 110 may present to the user 102 a guess as to what the user 102 is saying as the user 102 is saying it. Such a

guess may be refined repeatedly as the user 102 is speaking. In one implementation, if the audio recognition module 110 is guessing incorrectly, the user 102 may terminate the audio input 104 (by, e.g., releasing the button/key 108) and may again attempt to provide audio input 104. The audio recognition module 110 may interpret such a cancellation as an incorrect guess and may refine a learning model accordingly.

[0027] Once the recognized audio input 104 has been generated by the audio recognition module 110, the audio recognition module 110 may provide the recognized audio input 104 to the audio routing module 112. In some implementations, before providing the recognized audio input 104 to the audio routing module 112, the audio recognition module 110 may provide the recognized audio input 104 back to the user 102 to confirm that the recognized audio input 104 matches the audio input 104. The recognized audio input 104 may be displayed to the user, or may be translated by the audio recognition module 110 from text to audio and may be played to the user 102 by the telecommunication device 106. In one implementation, further action may not proceed until the user 102 confirms that the recognized audio input 104 is correct.

[0028] In various implementations, the audio routing module 112 comprises processes and rules for evaluating the recognized audio input 104 to determine how the recognized audio input 104 should be handled. The rules and processes comprising the audio routing module 112 may be implemented entirely or partially on the telecommunication device 106, or may be implemented on another device, such as a server of a telecommunication service provider. The audio routing module 112 may be implemented as a state machine comprising a finite number of states, each state associated with a different mechanism for handling the recognized audio input 104. Such states could be associated with mechanisms for handling the recognized audio input 104 as a request 114 made of a telecommunication device application or process, as an audio command 116, and/or as a network search request 118.

[0029] The audio routing module 112 could determine the state/mechanism based on a set of rules and/or priorities, such as the configurable rules or priorities added or modified by the rules or priorities process 126 (described further below) or by a carrier. For example, the set of rules may include a first rule that attempts to treat the recognized audio input as an audio command 116. In accordance with this rule, the audio routing module 112 may determine whether the recognized audio input 104 includes any known commands, such as the commands stored in the terms or macros data store 128 (described further below). If the audio routing module 112 determines that the recognized audio input 104 includes one or more of the commands, then the audio routing module 112 may handle the recognized audio input 104 as an audio command 116. If not, then the audio routing module 112 may next attempt to handle the recognized audio input 104 as a request 114 made of a telecommunication device application or process or as a network search request 118, depending on priorities specified by the set of rules. In one implementation, the rules may give priority to first attempting to handle the recognized audio input as a request 114 made of a telecommunication device application or process. In such an implementation, the audio routing module 112 may determine whether the recognized audio input 104 is associated with any sort of macros or data indicating an association between specific inputs and an application or process. For example, a weather application may be associated with a set of terms or phrases and/or may include a macro for recognizing certain terms or phrases and invoking the weather application. Such macros and terms or phrases may also be stored in the terms or macros data store 128. If the audio routing module 112 does not find a match or application/process capable of handling the recognized audio input 104, then the recognized audio input 104 may be treated as a network search request 118 and sent, for example, to a search service provider, such as Google® or Yahoo®.

[0030] In another implementation, the rules may give priority to first attempting to handle the recognized audio input as a network search request 118. The audio routing module 112 may submit the recognized audio input 104 as a query to search service. If the search results indicate, however, that no matches were found for the recognized audio input 104, then the audio routing module 112 may handle the recognized audio input as a request 114 made of a telecommunication device application or process. As described above, handling the recognized audio input 104 as a request 114 may involve comparing the recognized audio input 104 to terms or phrases associated with an application or to macros for interacting with an application, such as macros for invoking and passing the recognized audio input 104 to the application. Such terms or macros may be stored in the terms or macros data store 128.

[0031] In some implementations, the audio routing module 112 may determine how to handle the audio input 104 based on the interaction of the user 102 with the button/key 108. For example, if the user 102 presses and holds the button/key 108 for a predetermined amount of time (e.g., 3 seconds), the audio routing module 112 may handle audio input 104 as a query made of telecommunication device applications 114. If the user 102 presses and quickly releases the button/key 108, however, the audio routing module 112 may handle the same audio input 104 as a network search request 118. To enable this differential handling based on button/key 108 pressing behavior, different pressing behaviors may be associated with different rules or priorities.

[0032] In further implementations, the audio routing module 112 or an associated process may gather context information about the telecommunication device, such as location information, date/time, contacts, call log, browsing history, application execution context, etc., and perform the evaluating based at least in part on the context information. For example, if the user 102 is engaged in a phone call with

another user and, while engaged in the phone call, provides as recognized audio input 104 the command “send calendar”, the audio routing module 112 may determine that the user 102 wishes to send his or her calendar to the other user based on the fact that the user is engaged in a phone call with the other user. As mentioned above, to provide audio input 104 while engaged in another activity, including an activity that also involves providing audio input, the user 102 may press the button or key 108, which will cause the audio input 104 received during or after the press to be treated as the audio input 104 that is to be evaluated by the audio routing module 112. In other implementations, the execution context may indicate that the user 102 is involved in a chat session, an instant messaging session, or an email message exchange by writing the communication, speaking the communication, reading the communication, and/or listening to the communication. Such other execution contexts may each be taken into account by the audio routing module 112 in determining how to handle the recognized audio input 104. In one implementation, the context may provide a default that may be overridden by the contents of the recognized audio input 104. For example, if the user 102 were to say “send calendar”, the audio routing module 112 may interpret the recognized audio input 104 as a command 116 to send the calendar to the other user with whom the user 102 is currently communicating. If, on the other hand, the user 102 specifies a name in the audio input 104, such as the name of a contact of the user 102 stored in a contact list, the audio routing module 112 may ignore the context. Thus, if the user 102 were to say “send calendar to Eric Smith” while engaging in a phone call with Betty Crocker, the audio routing module 112 may send the calendar to Eric Smith rather than Betty Crocker. In one implementation, the audio routing module 112 may be configured (by rules or priorities) to, for example, send the calendar to both the name specified and the current communication partner.

[0033] In some implementations, while engaged in communication with another user and while providing audio input 104, the audio input 104 may, by default, not be relayed to the other user. In such embodiments, the default may be associated with a rule or priority which may be configured by user 102 or a carrier. Thus, for example, the user 102 may change the rule or priority to provide the audio input 104 by default. Also, in one implementation, words such as “private” or “public”, when included in the audio input 104, may be used to override a default.

[0034] In a further example, if the gathered context information is location information and the audio routing module 112 determines that the recognized audio input 104 should be handled as a network search request 118, the audio routing module 112 may include the location information as part of the network search request 118.

[0035] In some implementations, the audio commands 116 may include both commands executed on the telecommunication device 106 and commands executed on a remote server. Such commands 116 may include commands related to telecommunication device settings, telecommunication device personalizations, and/or on telecommunication device applications, such as applications, OS, etc. 120. For example, the audio routing module 112 may determine that the recognized audio input 104 matches a command associated with an airplane mode setting of the telecommunication device 106, and the audio routing module 112 may handle the command 116 by setting the airplane mode according to the recognized audio input 104 (e.g., turn airplane mode on or off). The audio routing module 112 may be configured to execute the command 116 or invoke an application or process to execute the command 116.

[0036] In another example, the audio routing module 112 may determine that the recognized audio input 104 matches a command associated with a query of a service provider for account information. Upon making the determination, the audio

routing module 112 may provide the recognized audio input 104 to the service provider, which may comprise a remote server, as a query for account information.

[0037] In further implementations, the recognized audio input 104 may be handled as a command 116 that is simultaneously executed on multiple applications or settings. If the audio routing module 112 determines that the recognized audio input 104 matches terms or phrases associated with multiple applications, then the audio routing module 112 may handle the recognized audio input as a command 116 to be executed on each of those applications or settings. For example, if the recognized audio input 104 is determined to include the words “turn on away message”, which match phrases associated with each of an email application and a text messaging application, then the audio routing module 112 may handle the input by executing (or invoking the applications to execute) the command 116 on both applications simultaneously. In one implementation, whether the recognized audio input 104 is executed on both applications simultaneously or only on one prioritized application may be determined by the user-configurable rules and priorities discussed above and below.

[0038] In some implementations, the recognized audio input 104 may be handled as a set of audio commands 116 associated with a macro. For example, if the recognized audio input 104 is “tell my wife that I will be home late for dinner”, the audio routing module 112 may locate a macro associated with the term “tell” which launches an email, call, or message with the contents “home late for dinner”. Those contents are then communicated to the person named in the recognized audio input 104.

[0039] In various implementations, the applications, OS, etc. 120 comprises any executing instructions on the telecommunication device 106. Such instructions include, for example, an OS of the telecommunication device 106, drivers for hardware

components of the telecommunication device 106, applications providing interfaces to settings or personalizations of the telecommunication device 106, applications made specifically for the telecommunication device 106, and third party applications. Collectively these applications/processes are hereinafter referred to as applications, OS, etc. 120. Such applications, OS, etc. 120 may be entirely or partially implemented on the telecommunication device 106. In some implementations, the applications, OS, etc. 120 may be implemented partially on another device or server.

[0040] In some implementations, as mentioned above, the applications, OS, etc. 120 may be invoked by the audio routing module 120 to execute a command 116 or a request 114. Such commands 116 may ask the applications, OS, etc. 120 to perform some task, such as initiate a call to a person specified in the recognized audio input 104 or change a setting of an application, OS, etc. 120. The requests 114 may ask the applications, OS, etc. 120 to return some data, such as the weather at a specified location, or pictures taken on a specified date. To execute both the commands 116 and requests 114, the audio routing module 112 must initialize an executing instance of the application, OS, etc. 120 – or discover such an instance – and invoke a function of the application, OS, etc. for receiving and processing recognized audio input 104 and associated commands 116 and requests 114. Such recognized audio input 104 and an indication of the sort of request 114 or command 116 may be passed as parameters to the function of the application, OS, etc. 120. To initialize and invoke the application, OS, etc. 120, the audio routing module 112 may make use of telecommunication-device-level macros stored in the terms or macros data store 128. For the OS and for applications 120 that come loaded on the telecommunication device 106 at the time of its sale to the user 102, the macros may also be provided and stored. For third party applications 120 installed by the user 102 on the telecommunication device 106, the macros may be received and stored at the time of installation, or at some later time. In

some implementations, described further below, macros may be built for the telecommunication device 106 by application developers 132 of third party applications 120 using the interfaces provided by the API 130 associated with the telecommunication device 106.

[0041] In further implementations, the application, OS, etc. 120 invoked by the audio routing module 112 may also be being used by the user 102 to enable communication with another user of another device. For example, the application, OS, etc. may be a telephony application 120 providing a local endpoint of a phone call. In such an example, the received command 116 may be a command 116 to conference in a third party. In response to receiving the command 116, the telephony application 120 may call and conference in the third party. In other implementations, the user 102 may be interacting with one application, OS, etc. 120, and the audio routing module 112 may invoke a different application, OS, etc. 120. For example, the user 102 may be using the telephony application 120 to engage in a phone call with another user, and the audio routing module 112 may initialize and invoke an email application 120 to, for example, email a calendar to another user, such as the other user conversing with the user 102 through the phone call.

[0042] In various embodiments, rather than executing a command or request contained in the recognized audio input 104 locally, the audio routing module 112 may send the determined command 116 or network search request 118 to a remote computing device via a networking fabric 122. The networking fabric 122 represents any one or more networks known in the art, such as cellular networks and/or data networks, including wide area networks (WANs), local area networks (LANs), personal area networks (PANs), and/or the Internet. For example, telecommunication device 106 could be connected to both cellular and data networks and the remote computing device could be connected exclusively to data networks. A connection

between one of the telecommunication devices 106 and the remote computing device may be through a number of routers, base stations, and/or telecommunication devices acting as bridges between cellular and data networks. Communications between the remote computing device and telecommunication device 106 utilize any sort of communication protocol known in the art for sending and receiving messages, such as the Transmission Control Protocol/Internet Protocol (TCP/IP) and/or the Hypertext Transfer Protocol (HTTP).

[0043] The remote computing device, in some implementations, may be associated with a search engine of a search service provider. Such a search engine may receive the recognized audio input 104 in the form of network search request 118 and may handle the network search request 118 as a search query. The search engine may then provide the search results to the telecommunication device 106 via the networking fabric 122. In other implementations, the remote computing device may be associated with the telecommunication service provider and may provide account information in response to receiving the recognized audio input 104 as a command/request for the account information. In yet other implementations, the remote computing device may be any sort of computing device known in the art, such as a web server, database server, or any other sort of server or device and may have any sort of application or data known in the art which may be the object of a command 116 or request 118 by the audio routing module 112 of the telecommunication device 106.

[0044] In various implementations, the results 124 of the execution of the request 114, command 116, or network search request 118 may be received by the telecommunication device 106 from the application, OS, etc. 120 or from a remote computing device via the networking fabric 122. The results 124 may then be provided to the user 102 in some form, such as via a UI or audio output. If the

recognized audio input 104 was handled as a command 116 by the audio routing module 112, the results 124 may simply comprise an indication that the command 116 was executed successfully or an error message indicating that the command 116 did not execute successfully. In one implementation, no results 124 may be provided in response to execution of a command 116. If the recognized audio input was handled as a request 114/118, then the information that was the object of the request 114/118 may comprise the results 124 that are provided to the user 102. In one implementation, if the user is engaged in communication with another user of another device, the results 124 may be provided to both users. Whether the results 124 are provided only to the user 102 or to the user 102 and other users may be determined with reference to the rules or priorities set by the rules or priorities process 126.

[0045] In various embodiments, as mentioned above, the audio routing module 112 may perform the evaluation of recognized audio input 104 based on configurable rules or priorities. Of those configurable rules or priorities, the user-configurable rules or priorities may be accessed, modified, added, or deleted by a rules or priorities process 126. The rules or priorities process 126 may be implemented on the telecommunication device 106, on another remote device or devices, or partially on both. In some implementations, the rules or priorities process 126 may include an audio and/or visual interface enabling users to have read and write access to the rules and priorities utilized by the audio routing module 112. In one implementation, the user 102 may be given only read access to some portion of the rules or priorities, or no access at all. The user 102 may add to, modify, or delete any one or more of the rules or priorities by providing audio input, text entry, and/or by selection of one or more UI controls.

[0046] In some embodiments, the rules or priorities may include one or more of a priority for attempting to handle the recognized audio input 104 as a command 116

before handling it as a request 114/118, a priority for attempting to handle the recognized audio input 104 as a request 114 of a telecommunication device application 120 before handling it as a network search request 118, or a priority for attempting to handle the recognized audio input 104 as a network search request 118 before handling it as a request 114 of a telecommunication device application 120. Additional rules or priorities may include a priority for one search service over another, a priority for simultaneous submission of a command 116 or request 114/118 to multiple applications 120 or remote devices, a priority for one application 120 over another if the command 116 or request 114 could be executed by either, or a priority for providing results 124 to only the user 102 or to both the user 102 and other user(s) in communication with the user 102. Also, any other sort of rules or priorities which may prioritize one course of action over another may be utilized. In some implementations, the rules or priorities may be stored in a rules or priorities data store (not shown).

[0047] In various implementations, as mentioned above, the audio routing module 112 may also or instead perform the evaluation of recognized audio input 104 based on commands or macros stored in a terms or macros data store 128. The terms or macros data store may, in some embodiments, comprise a database, such as a relational database, a structured database, an unstructured database, or an unstructured file. In addition to storing the terms or macros, the terms or macros data store 128 may also store metadata associated with the terms or macros. For example, the data store 128 may store metadata identifying the application associated with a given macro. Such a terms or macros data store 128 may be implemented locally on the telecommunication device 106 or remotely on another device, such as a server of a telecommunication service provider, or may be implemented in part locally and in part remotely. Further, the terms or macros data store 128 may respond to requests from

the audio routing module 112 by providing the contents sought by the audio routing module 112.

[0048] In some embodiments, the terms stored in the terms or macros data store 128 may include terms or phrases recognized as commands or requests associated with telecommunication device settings, personalizations, or applications (e.g., local or remote applications). The terms or phrases may come pre-loaded with the telecommunication device 106 when it is purchased or may be provided with a third party application at the time of installation or some later time. In one implantation, the telecommunication device 106 may provide an interface which allows a user to add to, modify, or delete terms or phrases in the terms or macros data store 128.

[0049] Also, in some implementations, the macros may comprise executable instructions for initializing an application, such as an application 120, for invoking the application, and for passing the recognized audio input 104 and other information to the application as parameters. The macros may also come pre-loaded with the telecommunication device 106 when it is purchased or may be provided with a third party application at the time of installation or some later time.

[0050] In further implementations, the rules or priorities process 126 and/or the terms or macros data store 128 may be accessible via a user interface. Such a user interface may enable the user 102 to see a list of commands, to edit that list, to see rules or priorities, and to edit those rules or priorities. The user interface could also be associated with the audio routing module 112 and/or the audio recognition module 110 and could, for instance, display a list of frequent queries. In one implementation, such a user interface may be associated with an icon, such as an icon on a main menu of the telecommunication device 106 which, when clicked on, pressed, or actuated may launch the user interface.

[0051] In various implementations, the terms, phrases, or macros may be registered with the telecommunication device 106 via an API 130 associated with the audio routing module 112. The API 130 may comprise a set of interfaces, such as a set of callable functions, which may be utilized by application developers 132 in building applications 120 for the telecommunication device 106 and other such devices. The callable functions may include functions for registering terms or phrases that are to be recognized as command or requests associated with the application 120, functions for notifying the audio routing module 112 of the ability of the application 120 to receive commands or requests, and functions for requesting more information from the audio routing module 112. Any other number of functions useful for making available the functionalities of the application 120 and for enabling communication between the audio routing module 112 and the application 120 may also be provided by API 130. In some implementations, the macros, applications 120, and/or installer application(s) associated with the applications 120 may be capable of calling the functions of API 130. The API 130 may be implemented entirely or partially on the telecommunication device 106.

[0052] FIG. 2 illustrates a network view showing a telecommunication device in communication with a remote server having audio recognition capabilities including at least a speech recognition component. As shown, a telecommunication device 106 may be in communication with a remote server, such as remote server 202, via the networking fabric 122. Both the telecommunication device 106 and networking fabric 122 are described above in greater detail above. As further shown, the remote server 202 may implement the audio recognition module 110 and a remote storage 204.

[0053] In various implementations, the remote server 202 may be any sort of computing device known in the art, except for the audio recognition module 110 and remote storage 204. The remote server 202 may be a personal computer (PC), a laptop

computer, a workstation, a server system, a mainframe, a telecommunication device, a personal digital assistant (PDA), a set-top box, or any other computing device known in the art. In other implementations (not shown), the remote server 202 may represent a plurality of computing devices each implementing some portion of the functionality represented by the audio recognition module 110 and remote storage 204. For example, remote server 202 could include a database server storing the remote storage 204 and a web server implementing the audio recognition module 110, as well as other devices. In one implementation, remote server 202 may refer to a first server device associated with a carrier (telecommunication service provider) and a second server device associated with an audio recognition service provider. In such an embodiment, the first server device may store the remote storage 204 and communicate with both the telecommunication device 106 and the second server device, and the second server device may include the audio recognition module 110 and communicate with the first server device, receiving the audio input 104 through the first server device.

[0054] In some implementations, the remote server 202 may be a server associated with an audio recognition service or with the telecommunication provider service. The remote server 202 may, in an implementation, receive audio input from the telecommunication device 106, such as audio input 104, and, utilizing the audio recognition module 110, may perform audio to text translation of the audio input, recognizing the audio input. The operations of the audio recognition module 110 are described above in greater detail. In one implementation, if the audio input is ambiguous, the remote server 202 may present optional meanings of the audio input to the user of the telecommunication device 106, either through an interface of the remote server 202 provided over the networking fabric 122 or through an interface provided by the telecommunication device 106.

[0055] In performing the speech recognition, the remote server 202 may rely on data stored in a remote storage 204, such as the learning models described above. Such a remote storage 204 may, in some embodiments, comprise a database, such as a relational database, a structured database, an unstructured database, or an unstructured file. The remote storage 204 may store data/learning model(s) associated with login identifiers, with service accounts, or with telecommunication devices.

[0056] FIG. 3 illustrates a component level view of a telecommunication device, in accordance with various embodiments. As shown, the telecommunication device 106 may include a memory 302, the memory storing modules/processes 304, applications, OS, etc. 120, and data 306. The telecommunication device 106 may further include processor(s) 308, a power supply 310, a display 312, a transceiver 314, sensors 316, output device(s) 318, input devices 320, and drive unit 322. Also, as shown, the input devices may include a button or key 108, and the drive unit 322 may include a machine readable medium 324.

[0057] In various implementations, memory 302 may generally include both volatile memory and non-volatile memory (e.g., RAM, ROM, Flash Memory, miniature hard drive, memory card, or the like). Additionally, in some implementations, memory 302 may include a SIM (subscriber identity module) card, which is a removable memory card used to identify a user of the telecommunication device 106 to a telecommunication service provider. The modules/processes 304 of memory 302 may, in some embodiments, include modules and/or processes such as one or more of the above-described audio capture process, the audio recognition module 110, the audio routing module 112, the rules or priorities process, process(es) associated with API 130, and/or any other modules or processes known in the art. Also, as shown, memory 302 may store the above described applications, OS, etc. Further, memory 302 may store data 306. Data 306 may, in some embodiments,

include data such as one or more of the above-described user-configurable rules or priorities, the terms or macros data store 128, and/or any other modules or processes known in the art.

[0058] In some implementations, the processor(s) 308 may be a central processing unit (CPU), a graphics processing unit (GPU), or both CPU and GPU, or other processing unit or component known in the art.

[0059] In further implementations, the power supply 310 may be implemented as one or more batteries. The power supply 310 might further include an external power source, such as an AC adapter or a powered docking cradle for supplementing or recharging the batteries.

[0060] In various implementations, the display 312 may be a liquid crystal display or any other type of display commonly used in telecommunication devices. For example, display 312 may be a touch-sensitive touch screen, and can then also act as an input device or keypad, such as for providing a soft-key keyboard, navigation buttons, or the like, as well as key 108 as a GUI key. FIG. 6 illustrates such a touch-sensitive touch screen in further detail.

[0061] In some implementations, the transceiver(s) 314 may include any sort of transceivers known in the art. For example, transceiver(s) 314 may include a radio transceiver and interface that performs the function of transmitting and receiving radio frequency communications via an antenna. The radio interface facilitates wireless connectivity between the telecommunication device 106 and various cell towers, base stations and/or access points. Transmissions to and from the radio interface may be conducted under control of the OS 120. For example, communications received by the radio interface may be disseminated to application programs 304/120 via the operating system 120, and vice versa

[0062] Transceiver(s) 314 may also include a near field interface that performs a function of transmitting and receiving near field radio communications via a near field antenna. For example, the near field interface may be used for functions, as is known in the art, such as communicating directly with nearby devices that are also, for instance, Bluetooth® or RFID enabled. A reader/interrogator may be incorporated into telecommunication device 106.

[0063] Also, transceiver(s) 314 may include a wireless LAN interface that performs the function of transmitting and receiving wireless communications using, for example, the IEEE 802.11, 802.16 and/or 802.20 standards. For example, the telecommunication device 106 can use a Wi-Fi interface to communicate directly with a nearby wireless access point such as for accessing the Internet directly without having to perform the access through the telecommunication service provider's network.

[0064] In various embodiments, sensors 316 may include any sort of sensors known in the art. For example, sensors 316 may include a global positioning system (GPS) transponder for endowing telecommunication device 106 with various GPS capabilities such as navigation, mapping, or the like. Further, telecommunication device 106 may include a motion sensor, such as an accelerometer for determining position, altitude and/or movement of telecommunication device 106, such as for adjusting display orientation, or the like.

[0065] In some implementations, the output device(s) 318 may include any sort of output devices known in the art, such as a display (already described as display 312), speakers, or vibrating mechanism, or tactile feedback mechanisms. Output device(s) 318 may also include ports for one or more peripheral devices, such as headphones, peripheral speakers, or a peripheral display.

[0066] In various implementations, input devices 320 may include the button or key 108 and any sort of input devices known in the art. For example, input devices may include a microphone (such as a microphone for receiving audio input 104), a keyboard/keypad, or a touch-sensitive display (such as the touch-sensitive touch screen described above). A keyboard/keypad may be a push button numeric dialing pad (such as on a typical telecommunication device), a multi-key keyboard (such as a conventional QWERTY keyboard), or one or more other types of keys or buttons, and may also include a joystick-like controller and/or designated navigation buttons, or the like. Button/key 108 has already been described above in great detail and, accordingly, will not be described further herein.

[0067] The machine readable medium 324 stores one or more sets of instructions (e.g., software) embodying any one or more of the methodologies or functions described herein. The instructions may also reside, completely or at least partially, within the memory 302 and within the processor 308 during execution thereof by the telecommunication device 106. The memory 302 and the processor 308 also may constitute machine readable media 324.

[0068] FIG. 4 illustrates a flowchart showing a method of receiving and handling audio input using a telecommunication device. As shown, a telecommunication device, such as telecommunication device 106, may receive an indication that a dedicated key, such as button or key 108, of the unified audio interface has been pressed or actuated, block 402, the dedicated key being reserved for invoking an audio capture process. In one implementation, the dedicated key may be a GUI key, such as GUI key 602, available on each UI screen rendered by the telecommunication device. In response to receiving the indication, the telecommunication device may then invoke the audio capture process, block 404. Next, the telecommunication device may receive the captured audio input, block 406.

[0069] The telecommunication device may then process, by an audio recognition module such as audio recognition module 110, the audio input to translate the audio input from audio to text, block 408. The telecommunication device can also learn, by the audio recognition module, voice characteristics of a speaker to improve audio recognition, block 410. Such voice characteristics belong to a speaker profile that is stored on a server and associated with a user login and/or user service agreement. The translating can both be based on the learning and cause occurrence of the learning.

[0070] In some implementations, the meaning of the audio input may be ambiguous. In response to detecting ambiguity, the telecommunication device may provide potential meanings of the audio input to resolve the meaning of the audio input, block 412.

[0071] Once the audio input has been processed, the telecommunication device may evaluate the audio input to determine whether to handle the audio input as a command, as a request of a telecommunication device application, or as a network search request, block 414, the determining being based on one or more user-configurable rules or priorities. Such evaluating may be performed, for example, by audio routing module 112 and may determine whether to handle the audio input as a command 116, as a request 114 of a telecommunication device application, or as a network search request 118.

[0072] In determining how to handle the audio input, block 414, the telecommunication device may determine whether the audio input comprises one or more recognized command words or one or more words associated with telecommunication-device-level macros, such as the terms or macros stored in terms or macros data store 128.

[0073] Also, the user-configurable rules or priorities used in performing the determining, block 414, such as the rules or priorities that are manageable through the rules or priorities process 126, may include one or more of a priority for attempting to handle audio input as a command before handling it as a request, a priority for attempting to handle audio input as a request of a telecommunication device application before handling it as a network search request, or a priority for attempting to handle audio input as a network search request before handling it as a request of a telecommunication device application.

[0074] Before or while performing the determining, block 414, the telecommunication device may also gather context information about the telecommunication device, block 416. The telecommunication device may then perform the determining, block 414, based at least in part on the context information or may submit the context information with a command or with a request.

[0075] After determining how to handle the audio input, block 414, the telecommunication device may execute the determined command, request of the telecommunication device application, or network search request or may provide the determined command, request of the telecommunication device application, or network search request to a system or application for execution, block 418. The telecommunication device application may include applications and processes such as applications, OS, etc. 120, and the network search request may be provided to remote computing devices via a networking fabric such as networking fabric 122. Also, block 418 may include carrying out the commands on telecommunication device settings, telecommunication device personalizations, and/or on telecommunication device applications. Further, the commands may include commands that are executed simultaneously on multiple telecommunication device applications and queries of service providers for account information.

[0076] Following execution/carrying out, block 418, the telecommunication device may receive results of the execution or carrying out of the command, request of the telecommunication device application, or network search request, block 420. Such results may include results 124.

[0077] At any point during, preceding, or following the execution of operations 402-420, the telecommunication device may expose an API, such as API 130, that enables telecommunication device applications to register with and/or interact with an audio routing module of the telecommunication device which performs the determining, block 422. In some implementations, the API may enable function call-based interaction between the telecommunication device applications and audio routing module while the audio routing module is performing the determining. Also, the API may enable a telecommunication device application to provide recognized commands or macros for use by the audio routing module in performing the determining.

[0078] FIG. 5 illustrates a flowchart showing a method of receiving and handling audio input using a telecommunication device while the telecommunication device is engaged in communication with another device. As shown, a user of a telecommunication device, such as telecommunication device 106, may be engaged in communication with another user of another device, block 502. In some implementations, the communication with the other device may be one of a phone call, a chat session, an instant messaging session, or an email message and engaging in the communication, block 502, may include writing the communication, speaking the communication, reading the communication, and/or listening to the communication.

[0079] While engaged in the communication, block 502, the telecommunication device may receive an actuation of a dedicated key, such as button or key 108, associated with a unified audio interface for audio commands and requests, block 504.

The telecommunication device may then process audio input provided with the actuation, such as audio input 104, as an audio command or request, such as a command 116 or a request 114/118. After processing the audio input, the telecommunication device may evaluate the audio input to determine whether to handle the audio input as an audio command, as a request of a telecommunication device application, or as a network search request, block 508, the determining being based at least in part on a context or data associated with the communication with the other device. Next, the telecommunication device may provide results of the command or request to the telecommunication device and the other device, block 510.

[0080] FIG. 6 illustrates example user interfaces of a telecommunication device, each rendering a dedicated key, in accordance with various embodiments. As shown, the telecommunication device 106 is capable of rendering a plurality of UI displays, each including a dedicated GUI key 602. UI displays 604 and 606 are illustrative of the multiple different displays each including the GUI key 602.

[0081] In various implementations, the dedicated GUI key 602 may have the same functionality as the button or key 108 described above. The telecommunication device 106 may either interact with each application or process associated with UI displays 604 and 606 to have those applications render the GUI key, or may provide an overlay UI in which the GUI key 602 is the only visible portion of the overlay.

[0082] UI displays 604 and 606 can be any sort of UI displays rendered by any applications or processes of the telecommunication device 106. In FIG. 6, UI display 604 is apparently a smiley-face picture or wallpaper and UI display 606 is an icon for an email or messaging application.

[0083] Although the subject matter has been described in language specific to structural features and/or methodological acts, it is to be understood that the subject matter defined in the appended claims is not necessarily limited to the specific features

or acts described. Rather, the specific features and acts are disclosed as exemplary forms of implementing the claims.

CLAIMS

WE CLAIM:

1. A method implemented in a telecommunication device, the method comprising:

receiving audio input through a unified audio interface of the telecommunication device; and

determining whether to handle the audio input as a command, as a request of a telecommunication device application, or as a network search request, the determining being based on one or more user-configurable rules or priorities.

2. The method of claim 1, further comprising:

receiving an indication that a dedicated key of the unified audio interface has been pressed or actuated, the dedicated key being reserved for invoking an audio capture process; and

in response to receiving the indication, invoking an audio capture process.

3. The method of claim 2, wherein the dedicated key is a graphic user interface (GUI) key available on each user interface (UI) screen rendered by the telecommunication device.

4. The method of claim 1, further comprising processing, by an audio recognition module, the audio input to translate the audio input from audio to text.

5. The method of claim 4, further comprising learning, by the audio recognition module, voice characteristics of a speaker to improve audio recognition.

6. The method of claim 5, wherein the processing includes retrieving the audio characteristics, the voice characteristics belonging to a speaker profile that is stored on a server and associated with a user login and/or user service agreement.

7. The method of claim 1, wherein the user-configurable rules or priorities include one or more of a priority for attempting to handle audio input as a command before handling it as a request, a priority for attempting to handle audio input as a request of a telecommunication device application before handling it as a network search request, or a priority for attempting to handle audio input as a network search request before handling it as a request of a telecommunication device application.

8. The method of claim 1, wherein the determining includes determining whether the audio input comprises one or more recognized command words or one or more words associated with telecommunication-device-level macros.

9. The method of claim 1, further comprising carrying out the commands on telecommunication device settings, telecommunication device personalizations, and/or on telecommunication device applications.

10. The method of claim 9, wherein the carrying out further comprises simultaneously executing the commands on multiple telecommunication device applications.

11. The method of claim 9, wherein the carrying out further comprises querying a service provider for account information.

12. The method of claim 1, further comprising exposing an application programming interface (API) that enables telecommunication device applications to register with and/or interact with an audio routing module of the telecommunication device which performs the determining.

13. The method of claim 1, further comprising providing potential meanings of the audio input to resolve the meaning of the audio input when the meaning of the audio input is ambiguous.

14. The method of claim 1, further comprising:
gathering context information about the telecommunication device; and
performing the determining based at least in part on the context information or submitting the context information with a command or with a request.

15. The method of claim 1, further comprising:

executing the determined command, request of the telecommunication device application, or network search request or providing the determined command, request of the telecommunication device application, or network search request to a system or application for execution; and

receiving results of the execution of the command, request of the telecommunication device application, or network search request.

16. A telecommunication device comprising:

a processor;

a dedicated key configured to invoke an audio capture process when actuated;

an audio recognition module operable by the processor and associated with the dedicated key, the audio recognition module configured to recognize audio input captured by the audio capture process; and

an audio routing module operable by the processor and configured to evaluate the recognized audio input to determine whether to handle the recognized audio input as a command, as a request of a telecommunication device application, or as a network search request, the determining being based on one or more user-configurable rules or priorities, the audio routing module offering an application programming interface which enables telecommunication device applications to register with and/or interact with the audio routing module.

17. The telecommunication device of claim 17, wherein the application programming interface provided by the audio routing module enables a

telecommunication device application to provide recognized commands or macros for use by the audio routing module.

18. An article of manufacture comprising:

one or more storage media; and

executable instructions stored on the storage media which, when executed by a telecommunication device, instruct the telecommunication device to perform operations including:

receiving while engaged in communication with another device, an actuation of a dedicated key associated with a unified audio interface for audio commands and requests;

processing audio input provided with the actuation as an audio command or request; and

determining whether to handle the audio input as an audio command, as a request of a telecommunication device application, or as a network search request, the determining being based at least in part on a context or data associated with the communication with the other device.

19. The article of claim 18, wherein the communication with the other device is one of a phone call, a chat session, an instant messaging session, or an email message and engaging in the communication includes writing the communication, speaking the communication, reading the communication, and/or listening to the communication.

20. The article of claim 18, wherein the operations further include providing results of the command or request to both the telecommunication device and the other device.

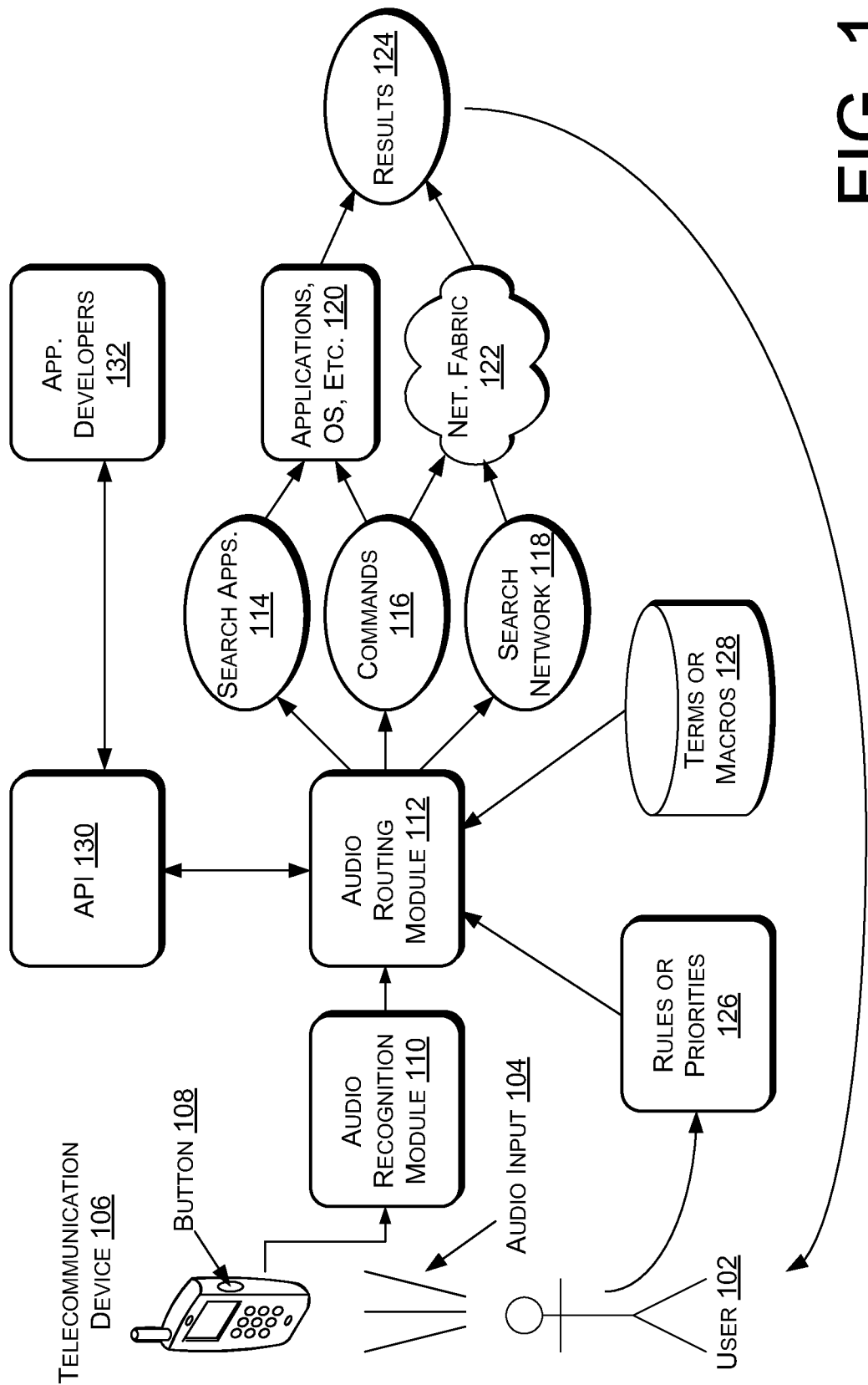


FIG. 1

2 / 6

+

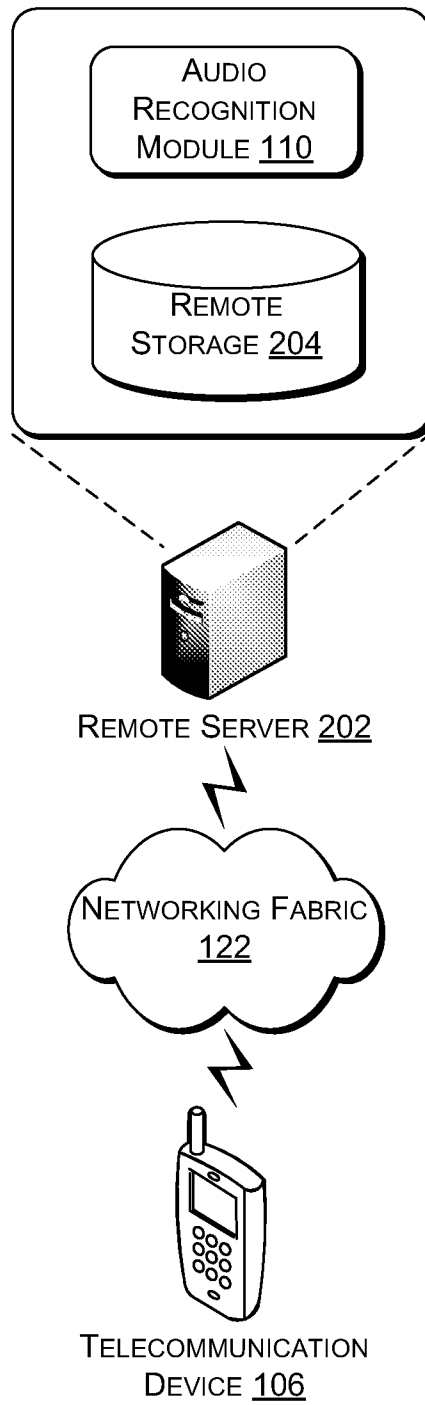


FIG. 2

+

3 / 6

+

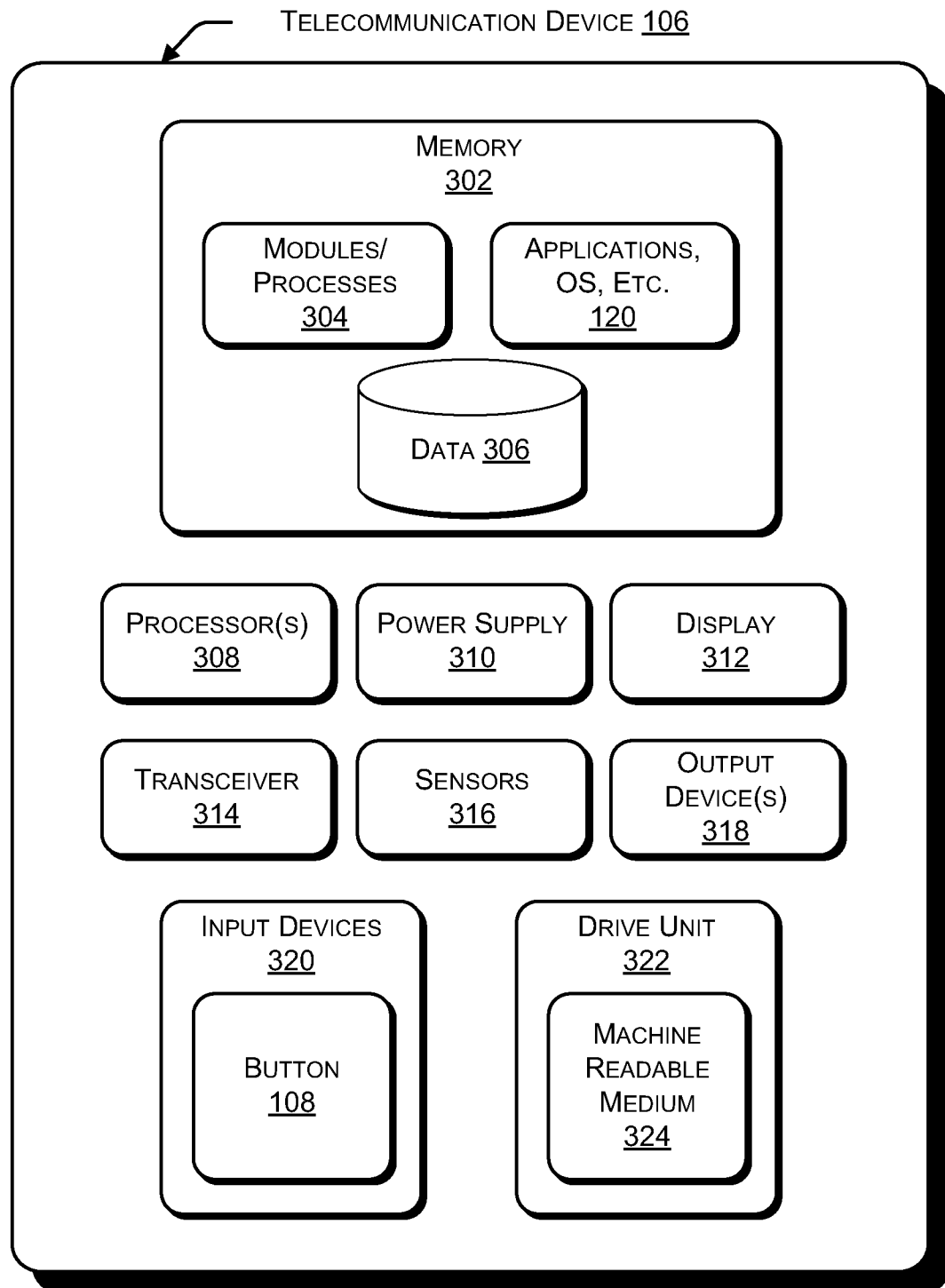


FIG. 3

+

4 / 6

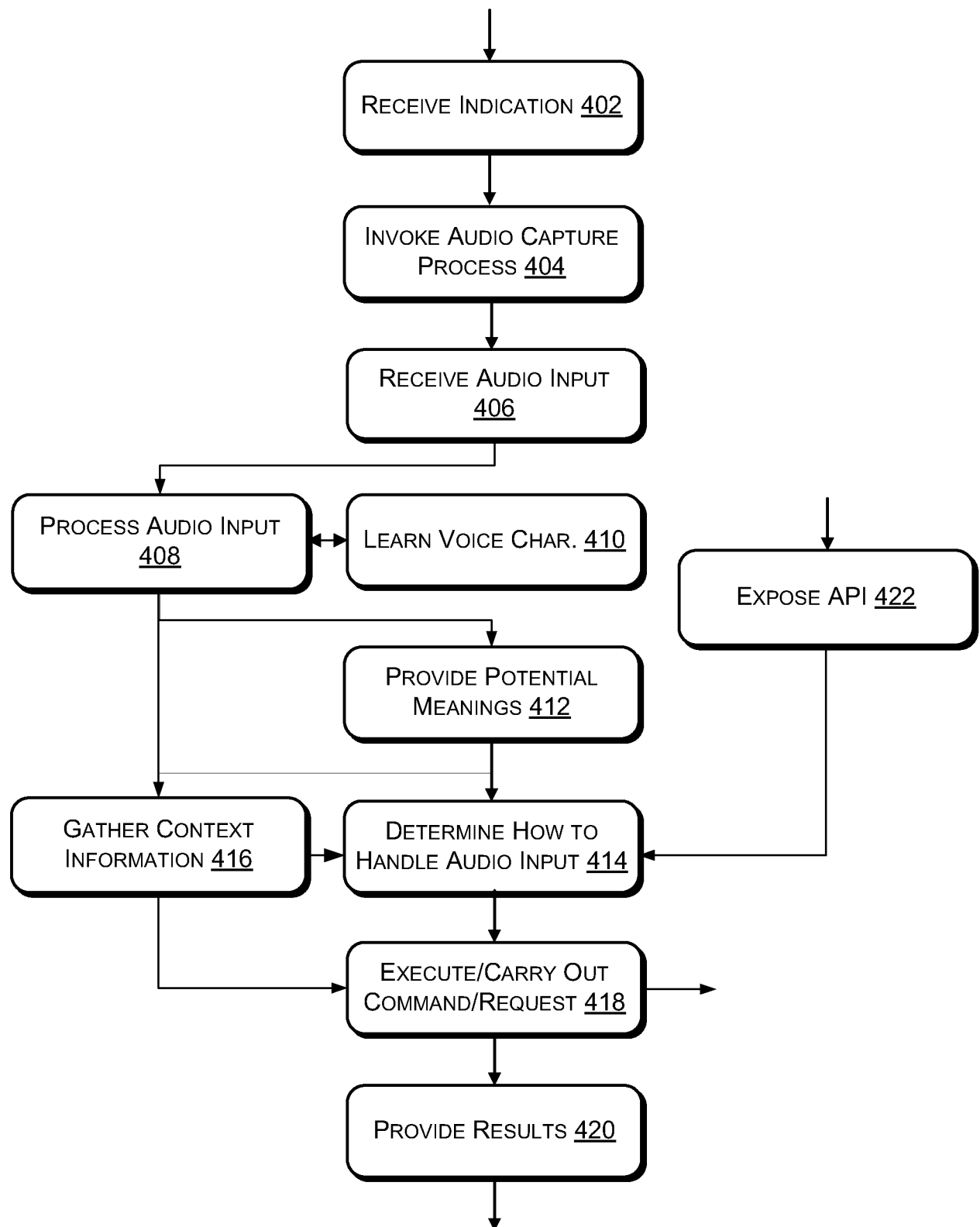


FIG. 4

5 / 6

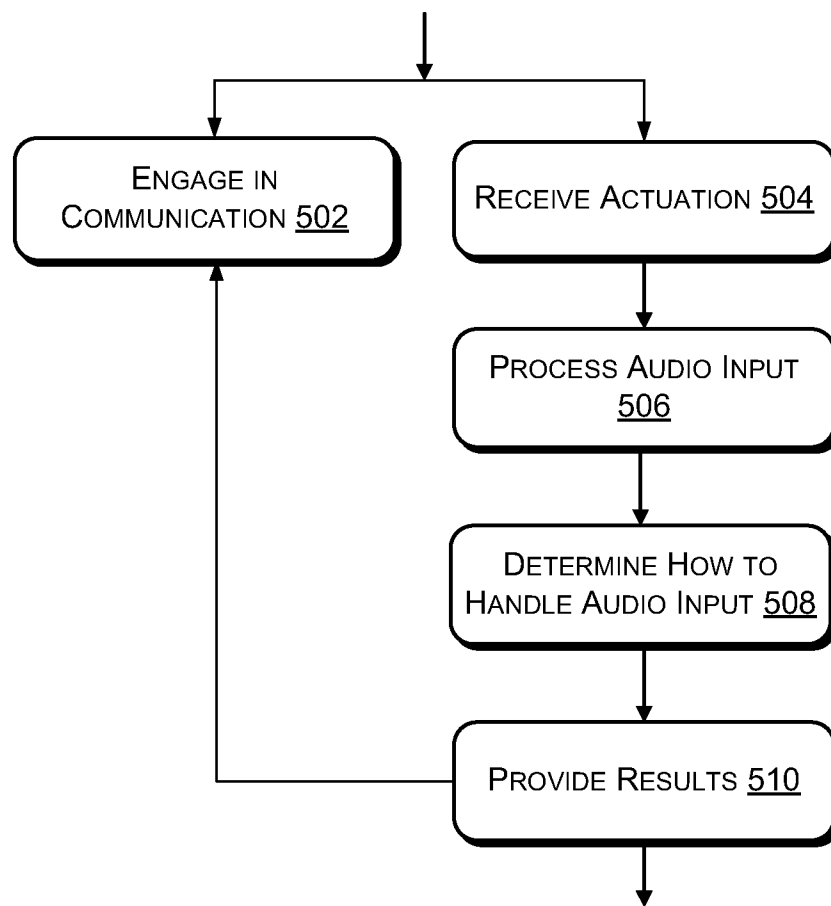


FIG. 5

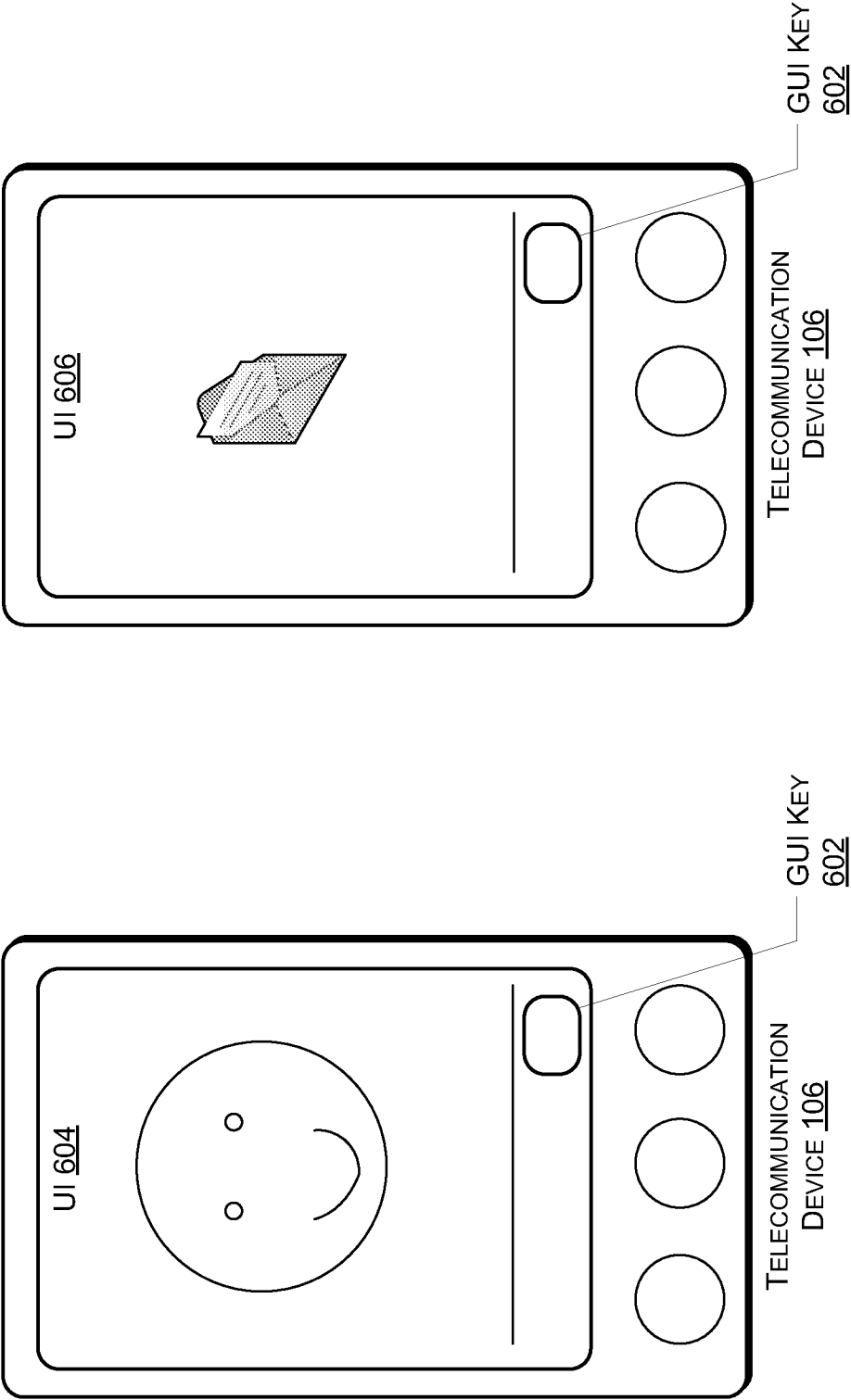


FIG. 6